An Exploration of Students' Construction of Meaning through Symbolic Manipulation and Table / Graph Use in Statistical Inference Tasks: The Cases of Normal and t Distributions

CHE-YU SHEN

Ph.D.

2012

An Exploration of Students' Construction of Meaning through Symbolic Manipulation and Table / Graph Use in Statistical Inference Tasks: The Cases of Normal and t Distributions

by

CHE-YU SHEN

A thesis submitted for the degree of Doctor of Philosophy in the

School of Education and Lifelong Learning,

University of East Anglia

March 2012

ACKNOWLEDGEMENT

This research would not have been possible without the help from many people. I would like to take this opportunity to express my sincerest gratitude to those people who helped me complete this journey. First, I give my greatest appreciation to my supervisors, Dr. Paola Iannone and Dr. Elena Nardi, for their patience, guidance and encouragement to lead me to finish the study. Second, I need to thank my parents for their financial and mental support. Third, I want to thank the Institute of Technology for sponsoring me and providing me with an opportunity to carry out the data collection, and the participating teacher and students for cooperating with me during the eleven weeks. Besides, I would like to thank my beloved wife Rong who always stays with me. Last but not least, I want to thank to all my friends and colleagues who have helped me in many ways.

ABSTRACT

This study investigates college students' use of statistics tables when solving problems on normal distribution and t distribution. Particular attention is given to the way in which students use the graphical representation of the normal curve and the t-curve in their solutions. A review of the literature on the teaching and learning of statistics at undergraduate level reveals that not much work has been carried out to investigate how students use statistics tables. The data in this study was collected in a business school at a private institute of technology in the south of Taiwan. Ten students in the second year of their course and their teacher participated in the study. The students were interviewed three times during the course of one semester. The data collected include field notes, audio recording and photos of classroom observation; participants' answer sheet in the mid-term and final examinations, and exercise questions and audio/video recordings in the interviews. The main body of data are the clinical interviews carried out with the students. In these interviews the students were asked to solve statistics problems using a talk-aloud technique. The interviews were audio recorded and fully transcribed. The interview data were analysed by decomposing the students' answer into the solving steps used in the solution of each problem. Analysis of the participants' solutions revealed that using the tables of distribution to find the solution to the given task was problematic. Their solution attempts can be categorised into six types, but the underlying difficulty appeared to be the symbolic manipulation of the data in the question. Students seem not to ascribe statistics meaning to the symbols and tend to perform symbolic manipulations without investigating the meaning of the symbols first. Moreover, most participants did not use graphs when they solved the problems, and only four participants actively used graphs in a few questions, perhaps to visualise the values in the questions or to create meaning. The students who consistently used graphs in their solutions on the whole performed better than the ones who didn't across the topics. The study concludes with some recommendations for the teaching of statistics as a service subject.

CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Emergence of the problem

Statistics learners may have varying difficulties in learning this subject. Some difficulties may be caused by carelessness or by poor arithmetic skills, such as misreading the questions or miscalculation. Other difficulties may be caused by more complex factors. For instance, some students' self-created theories are capable of explaining current situations, so they feel that they have understood even though their understanding may be inconsistent with the theory (Sierpinska, 1994). Other factors include lack of prior knowledge (such as being confused about the concepts of variance and standard deviation) or being unable to distinguish the condition of questions, etc. Teachers need to know the ideas with which students often have difficulties and find ways to help bridge common misunderstandings (NCTM, 2000).

I taught statistics for three and a half years (seven semesters) in a business school at a private Institute of Technology in Taiwan. In the statistics classes, I often found that students had difficulty using the tables of distribution correctly, no matter which distribution they encountered. Many of my students would go on to an optional course of statistics software (SPSS), but some of them had difficulty operating the software because they did not understand what the terminology meant, such as distinguishing the standard deviation from the standard error, or asking what the p-value is when doing the hypothesis testing. The idea of p-value, which is the left-tail or right-tail area below the curve of distribution, was used in the tables of distribution, but it seemed that some students failed to realise this. Batanero (2004) stresses that the idea of the level of significance, α, linked to the p-value is a "particularly misunderstood concept" (2004, No pagination (hereafter, "n. p.")) in hypothesis testing. Without the ability to use the tables of distribution, students are unable to learn well and will feel frustrated in learning statistical inference.

As a lecturer teaching a college-level statistics course, I used the table of normal distribution with a p-value in the left tail (see Appendix 1.1.) because the table included both positive and negative z scores. There is a graph above the left-half of the table and another graph above the right-half of the table. However, I found that some students were confused by the values inside/outside the table, in the formula and in the graphs, so I created a video to help the students understand the relationship between the table of normal distribution and the graph of normal curve with a left-tail (a Chinese version of the video is attached with the thesis or available at: http://youtu.be/Xb6HKuPbX0U).

The main feature of the video is that it looks at both the graph of the normal curve and the table of normal distribution at the same time. Three hundred and sixty z scores (including -0.00 and negative values) were included in the left-half of the table, and three hundred and sixty z scores (including 0.00 and positive values) were included in the right-half of the table. The video does not introduce every value in the table. Instead, it introduces a few values to explain in which place a z score or a probability belongs in the table, how to find individual z scores in the table, and the relationship between the elements in the table and the elements in the graph above the table. The video also compares and contrasts a few values which often confuse students.

At the beginning of the video, I scanned the original table used in my class (i.e. Appendix 1.1.), and then added an explanation of the meaning of the shadow area (see

Appendix 1.2.). I gave the first example to show how P(Z < -3.50) = 0.0002 was displayed in the table, by circling -3.5 on the left-hand side and .00 on the top of the left-half of the table, and underscoring the .0002 on the upper-left corner inside the table (see Appendix 1.3.).

However, the given graphs did not represent the correct location of the z score and the size of probability in this example. In fact, not all z scores and probabilities could be displayed in the attached graphs, and it was possible that the two fixed graphs did not help but rather disrupted students' thinking. Therefore, I drew another graph above the two tables with varying z scores and corresponding probabilities. I also encouraged students to draw graphs when they used the table to find values.

The varying graphs began with P(Z < -3.59) = 0.002 and finished at P(Z < 3.59) = 0.998. The video enabled students to see how the values moved in the table and in the graph at the same time. When the z score increased from -3.59 on the right-hand side of the first row, through –3.54 in the middle, toward -3.50 on the left-hand side, downward to -3.49 on the right-hand side of the second row, to -2.00 on the left-hand side of the sixteenth row, and then to -0.90 on the twenty-seventh row, viewers could observe the red left-tail gradually and continuously expanding with the z score moving forward to the right-hand side (see Appendix 1.4.). Some values which often confused students in the table, such as P(Z < -0.90) and P(Z < -0.09), were also compared in the table (see Appendix 1.5.).

The z scores -0.00 and 0.00 were the same, but -0.00 was on the left-hand side of the table and 0.00 was on the right-hand side of the table. This was also explained and linked in the video (see Appendix 1.6.).

On the right-hand side of the table, the positive z score moved forward to the right-hand side and downward, and the red left-tail continuously grew. When the z score came to 3.59 and the probability equalled 0.9998, the red left-tail occupied nearly the whole area below the curve, and the white right-tail became very small (see Appendix 1.7.). The video ends with a question (see Figure 1.1.).

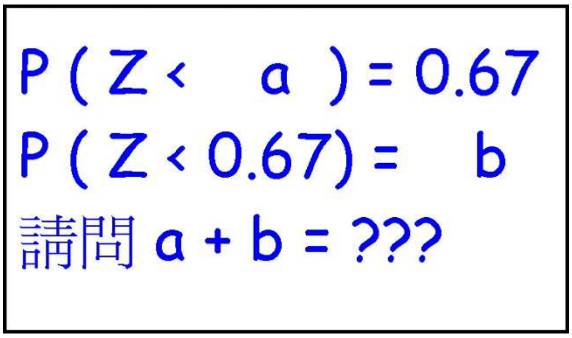


Figure 1.1. Question at the end of the video.

It was my hope that the students, after watching the video of the varying graphs, would have clearer ideas of what ‘a’ and ‘b’ meant respectively, where they were inside/outside the table, and how they were shown in the graphs. Some students told me that this video was very helpful for them to understand how to use the table of normal distribution, but I did not collect the students’ answers and was thus unable to know whether the video really made a difference. However, the video inspired me to further investigate how students learn to use the tables of distribution and the graph of the distribution curve, and what difficulties they may have.

I looked for suggestions to improve students’ learning of tables of distribution in the literature, but I did not find relevant research, so I contacted two researchers via email: J.M. Shaughnessy, who was the president of NCTM, and C. Batanero, who had been the president of IASE. They replied to me that, as far as they knew, there was scarce research on this topic. Therefore, I decided to make a contribution to research in this area.

Legutko (2008) points out that teachers often underestimate the necessity for students to master basic skills, such as data comprehension. According to the scarcity of relevant research and my own observation in my statistics classes, it seems that using the tables of distribution is regarded as a basic skill, yet it is given little attention by teachers and researchers. However,

> [w]hat seems simple and obvious in mathematics does not necessarily have to be simple and obvious in mathematics teaching (Legutko 2008, p.142).

Since there is little research focusing on this situation, and because some students have indeed had difficulties when using the tables of distribution, I would like to investigate this issue, in the hope that it may further our knowledge and inspire additional research. In the present study, I am going to investigate how students learn and what difficulties they may have when using the tables of distribution in the two parts of statistical inference, estimating and hypothesis testing.

Cosmides and Tooby (1996) find that active involvement with a pictorial representation helps to make necessary information explicit and boosts performance. Breen (1997) stresses the “therapeutic” (1997, p.97) use of mathematical and educational imageries as a “tool for aiding learning” (1997, p.97). In Brase’s (2009) study of improving Bayesian reasoning, a suitable pictorial representation was found to be helpful in learning statistics, but not all types of pictorial representations are equivalently useful. In fact, using visualisation does not guarantee the correct solution unless the visualisation is correctly drawn. Brase (2009) explains that many of the actively constructed visual forms are not correctly completed, and incorrect constructions may end up confusing students rather than helping them.

Some participants in the study argued that a big problem in learning statistics is that there are too many formulas for them to distinguish the condition of the descriptive question and select the corresponding formula(s). Therefore, most exercise questions were presented in symbolic forms in order to help them focus on the research questions.

1.2 Research questions

Statistics is a subject with numerous formulas and symbols. Some students are used to symbolic representations and reluctant to use visual methods for meaningful learning, even if particular images or diagrams can aid mathematical generalisation and link different modes of thinking (Healy and Hoyles, 1996), and "[r]epresenting instances can reveal the set structure of a problem" (Sloman *et al.*, 2003, p.298). Thus, this study not only explores how students use the tables of normal distribution and t distribution, but also investigates the role of visualisation in learning and problem-solving in statistics tables. Therefore, the research questions are:

1. How do the students use the tables of distribution to solve questions in statistical inference, and what difficulties do they have?
2. How does visualisation help students use the tables of distribution, and what difficulties do they have?
3. Why do some students not use visualisation to help in using the tables of distribution, and how do the students who use visualisation perform in comparison with those who do not use visualisation?

1.3 The organisation of the chapters

This chapter explains the background, purpose, focus and three research questions of the study. Here I will briefly introduce the content of the following chapters:

Chapter 2 reviews literature related to the study. This chapter includes two main parts, which are literature about statistics teaching and learning and literature about visual thinking.

Chapter 3 introduces the background of the participants and the site of data collection. It also explains the educational systems in Taiwan.

Chapter 4 explains the methods and methodology used in this study, including the methods of data collection and the process and tool of data analysis.

Chapter 5 introduces the content of the participants' statistics course, describes the students' use of the tables of normal distribution and t distribution, and compares different types of table of normal distribution and the transferring principles.

Chapter 6 explains the design of the clinical interviews. I will first discuss the findings from the class observations and the materials analysis, and then explain how these findings relate to the study and link the findings to the questions for interviews.

Chapters 7, 8, 9 and 10 offer analyses of participants' solutions in four kinds of exercise questions respectively by deconstructing their solving steps in the exercise questions. Chapter 7 analyses how the participants used the tables of normal distribution to find a value; Chapter 8 analyses how they used the table of t distribution to find a value;

Chapter 9 analyses how the participants used the tables of normal distribution to find the z score of confidence interval; and Chapter 10 analyses how they used the tables of normal distribution to find the critical value and critical region of hypothesis testing.

Chapter 11 discusses the influence, benefits and difficulties of the use of visualisation in the study, explains the reasons for participants' reluctance to use visualisation, and compares the performances of the participants who used graphs with those who did not use graphs in different kinds of question. This chapter also examines the participants' drawing of normal curves according to the symbolic form in order to explore the participants' interpretation of the relationship between mean, standard deviation and the curve.

Chapter 12 quickly reviews and summarises the exercise questions and the participants' solutions and explanations, which have been analysed in Chapters 7, 8, 9, 10 and 11, in order to help provide a clearer view of the performance and tendency of the participants' responses.

Chapter 13 replies to the three research question of the study with a discussion based on the data analysis in Chapters 7 to 12.

Chapter 14 offers reflections on this study and provides implications for future education and research according to the findings in this study.

CHAPTER 2

LITERATURE REVIEW

The literature review is based on research key words such as statistics education and visualization to search relevant journals and publications. However, research of learning tables of distribution was scarce. Therefore, I searched and linked papers relevant to the research questions including frameworks of interpreting questions, how students learn tables, and whether/how they use visual presentations in learning mathematical subjects. There will be two parts in this chapter.

In the first part of this chapter, I will introduce the development of research of statistics education in the past and present, some principles for learning statistics, theoretical frameworks for analysing data, challenges in statistics education and the teaching and learning of statistical inference. Then, in the second part, I will introduce the development of visual thinking, how visualisation helps with learning mathematical subjects and the factors of comprehending graphs. I will also discuss how number sense, symbol sense and graph sense relate to this study.

2.1 Statistics learning and teaching

In the first part of the literature review, I will briefly introduce the development of statistics education and the previous research on difficulties in and suggestions for teaching and learning statistical inference, which is the focus of this study. In addition, I will introduce the theoretical frameworks used to analyse data in statistical literacy, statistical reasoning, and statistical thinking. I will also include recent research on statistics education and its challenges,

2.1.1 Development of statistics education

Statistics and probability entered the standard curriculum of mathematics in European countries in the middle of the twentieth century, but it only entered the mathematics curriculum in the United States in 1989, which marks the first time that statistics was "placed… on an equal footing with symbol sense, algebra, geometry, and measurement as a critical foundation stone for school mathematics" (Shaughnessy, 2007, p.957). Recently, statistics has been widely taught in many universities as a tool for problem-solving in other fields, such as social science and business administration. However, the typical pedagogy of statistics, focusing on teaching formulas for calculating statistics, is "over-mathematized" (Artigue *et al.*, 2007, p. 1035) and ignores data context, interpretative activities or simulations which have proved helpful in building stochastic intuitions. Therefore, students without prior or concurrent mathematical knowledge may have difficulties in learning statistics.

Unlike education in other mathematical subjects such as calculus and algebra, statistics education is a relatively new field of enquiry and seems to be "invisible" and "fragmented" (Garfield and Ben-Zvi, 2008, p.21). This emerging discipline has a research base which is difficult to locate and build upon. Since the International Statistical Institute (ISI) established an education committee in 1949, conference proceedings such as The International Conferences on the Teaching Statistics (ICOTS), and journals such as Teaching Statistics (TS), Journal of Statistics Education (JSE) and Statistics Education Research Journal (SERJ) have become international platforms to share ideas on improving statistics education. Some studies related to statistics education are also published by Psychology of Mathematics Education (PME) and the International Congress on Mathematics Education meetings (ICME). Over forty-seven doctoral dissertations about statistics education have been published between 2000 and 2008 (Garfield and Ben-Zvi, 2008). The increasing volume of research on statistics education shows that this area has been noticed by researchers and educators.

Research on students' understanding of statistics and probability has been developed in Europe for several decades. The initial work in this field, led by Piaget and Inhelder (1951), focused on the "developmental growth and structure of people's probabilistic thinking and intuitions" (Garfield and Ben-Zvi, 2008, pp.22-23). During the 1970s, some psychologists investigated how people made judgements and decisions in uncertain situations. For instance, Kahneman *et al.* (1982) identify incorrect ways of reasoning. Their heuristics and biases programme explains people's frequent statistical thinking when they make decisions under uncertainty, namely representativeness and availability, and the systemic errors heuristics users may have. At that time, studies in this area focused on how people understood or misunderstood particular statistical ideas. For example, Falk (1988) identifies misconceptions in conditional probability, and Konold (1989, 1991) considers misconceptions in independence and randomness. Some people were found to possess a way of reasoning called outcome orientation which made them "make yes or no decisions about single events rather than looking at the series of events" (Konold, 1989, cited in Garfield and Ben-Zvi, 2008, p.24).

Subsequent studies focused on "methods of training individuals to reason more correctly" (Garfield and Ben-Zvi, 2008, p.24). Some researchers argue that many misconceptions are caused by people's inability to use proportional reasoning and that some of the difficulties in learning statistics may be improved by different ways of presentation. For example, Sedlmeier (1999) notices that the mathematical equivalency between the presentations does not guarantee psychological equivalency. Some learners cannot deal with probability value presented in the format of absolute frequencies as successfully as when they deal with rates or percentage, but training students to translate the statistical task to a proper format such as tree diagrams or absolute frequencies can help.

Therefore, Garfield and Ben-Zvi, (2008) suggest that learners use frequency approaches because they find that people learn better in using counts and ratios than in using percents and decimals. Some studies examined misconceptions in other fields, such as contingency tables (Batanero *et al.*, 1996) and significance tests (Falk and Greenbaum, 1995). Garfield (2003) suggests Statistical Reasoning Assessment (SRA) as a method to explore students' correct and incorrect reasoning in probability and statistics by letting the students choose the response closest to their thinking.

Instead of focusing on misconceptions or errors of particular statistical ideas, some researchers investigated how students begin to develop good reasoning and understanding of basic statistical concepts. These studies uncovered some difficulties students experience when they learn concepts which were thought basic or simple by educators and teachers, such as the mean (Shaughnessy, 2007).

In recent years, the field of statistics education has expanded, and researchers have started to distinguish between the research of understanding probability and the research of understanding statistics, as these are two separate disciplines. Research in statistics education includes students' knowledge and reasoning, and teachers' knowledge and teaching practice, but most have investigated particular ideas in students' understanding, such as centres, variation, comparison of data sets, learning with technology, and understanding of graphs. Shaughnessy (2007) points out that students' understanding of statistical inference is still an area which needs further investigation. I will discuss this in Section 2.1.2.

Many studies of statistics education investigated the teaching and learning of statistics at the college level. Some of the studies focused on a technological tool such as computers and graphing calculators or a teaching method such as simulation training or a collaborative classroom, while others “examined a particular activity or intervention” (Garfield and Ben-Zvi, 2008, p.28). For example, Quilici and Mayer (2002) find that more college students solve questions on the basis of structural features rather than surface features after they finish the statistics course, and Lovett (2001) suggests that feedback given to students helps them to select appropriate methods to solve problems.

Some studies investigate the non-cognitive factors influencing students’ success in statistics, such as students’ attitude and anxiety about statistics (Schau and Mattern, 1997), relationships between students’ characteristics, including mathematical background, statistics attitudes or statistics anxiety, and comparison of the outcome or understanding of students who take statistics courses in education, psychology or social science (Wisenbaker and Scott, 1997; Earley, 2001). Surprisingly, these studies did not show a positive correlation between students’ mathematics abilities or good attitudes and their performance in statistics. Garfield and Ben-Zvi (2008) suggest that “motivation, conscientiousness and desire to learn may be better predictors” (2008, p.33).

According to Garfield and Ben-Zvi (2008), it is difficult to determine the influence of a particular teaching method or instruction tool on students’ learning in a course because of the limitation of study design or assessment. They explain that the results of comparative research are limited to particular courses and cannot be generalised to other courses. Although this study focuses on students’ actual use of the tables of normal distribution and the table of *t* distribution and the influence of visualisation, I believe that its results can also help teaching and learning the table of chi-square distribution and the table of F distribution, such as to derive meaning by linking the elements in the tables and the elements in the graphs.

Garfield and Ben-Zvi (2008) notice the problematic situation that final exam scores and course grades are often used as outcome measures in many studies; this is problematic because of the “lack of high quality and consistent measures used to assess student learning outcomes” (2008, p.33), so they suggest three levels of understanding in learning statistics to distinguish learning outcomes, namely statistical literacy, statistical reasoning and statistical thinking.

First, statistical literacy is the expected ability gained by education in an information-laden society, and an important part of adults’ literacy and numeracy. Gal (2002) explains statistical literacy as the ability to interpret, to evaluate, and to communicate with statistical information in available data. Watson and Callingham (2003) add that statistical literacy can be broken down into three levels, which are knowledge of terms, understanding of terms in context, and critiquing of claims in the media. Garfield and Ben-Zvi (2008) explain that

> Statistical literacy involves understanding and using the basic language and tools of statistics: knowing what basic statistical terms mean, understanding the use of simple statistical symbols, and recognizing and being able to interpret different representations of data… statistical literacy provides the foundation for reasoning and thinking (Garfield and Ben-Zvi, 2008, pp.34-35).

This study will examine the participants’ ability to interpret and connect different representations of data (i.e. question in symbolic form, graphical form and table).

Second, statistical reasoning is the way in which people possess statistical ideas reason and how they make sense of available statistical information. It means not only understanding but also being able to explain statistical processes and interpret the results (Garfield, 2002). It is the “mental representations and connections that students have regarding statistical concepts” (Garfield and Ben-Zvi, 2008, p.34), and it involves connecting one concept to another, such as centre and spread, or the combining of ideas about data and chance.

Third, statistical thinking is how professional statisticians think. It is the “normative use of statistical models, methods, and applications in considering or solving statistical problems” (Garfield and Ben-Zvi, 2008, p.34). Therefore, it involves a higher order of thinking than statistical reasoning and statistical literacy. Garfield and Ben-Zvi (2008) explain that statistical thinking includes

> …knowing how and why to use a particular method, measure, design or statistical model; deep understanding of theories underlying statistical processes and methods; as well as understanding the constraints and limitations of statistics and statistical inference (2008, p.34).

Garfield and Ben-Zvi (2008) argue for the integration of learning outcomes and learning goals:

> Clarifying desired learning outcomes can also help researchers better develop and use appropriate measures in their studies, and to align these measures with learning goals valued by the statistics education community (Garfield and Ben-Zvi, 2008, p.33).

The emphasis of statistics instruction has moved from developing students' procedural understanding to developing their conceptual understanding and statistical literacy, reasoning and thinking, especially on the three big ideas of distribution, centre and variability. Garfield and Ben-Zvi (2008) notice some constraints in some previous studies, such as "[l]ack of connection to theory or previous research", "[l]ack of focused questions on the effects of a particular tool or method or activity on learning to understand a particular concept" and "[f]ailure to look beyond the researcher's own discipline to other related studies" (2008, p.36). They continue that the use of final exams as the measure of student learning in some studies may not accurately reflect students' conceptual learning results if the items in the exam can be simply computed or memorised, and that explanatory variables should be well defined and consistent in the study so that the results of the study can be useful and generalised to other educators. They find that some studies

> reveal the types of difficulties students have when learning particular topics, so that teachers [and educators] may not only be aware of where errors and misconceptions might occur and how students' statistical reasoning might develop, but also what to look for in their informal and formal assessments of their learning… the research literature is especially important to consider because it contradicts many informal or intuitive beliefs by teachers. (Garfield and Ben-Zvi, 2008, p.39)

Artigue *et al.* (2007) suggest two important issues in the teaching of statistics: "[m]odeling and problem solving" (2007, p.1035). They explain that statistical problems have multiple possible solutions since the questions are always "open-ended or ill-defined" (2007, p.1035), so selecting the model is more difficult than the following mathematical reasoning. Even if the context of a statistics problem can guide the statistical model selection, the context can also cause reasoning errors.

Kahneman *et al.*'s (1982) heuristics and biases theory has inspired researchers (such as Shaughnessy, 1992) to investigate misconceptions, overconfidence in the results of statistical tests, and underestimating the width of confidence intervals.

Cited in Artigue *et al.* (2007), Estepa (1993) finds that mathematical concepts may cause difficulties when learning correlation, such as correlation coefficient r1 = -0.7 regarded as a smaller degree of dependence than r2 = 0.1 simply because -0.7 < 0. However, such misconceptions caused by numbers can be improved by different representations such as graphs. Batanero *et al.* (1996) report that students perform badly in establishing judgment about association when they read contingency tables. They explain the situation named "illusory correlation", where students hold expectations and beliefs about the relationship between the variables and ignore the evidence of independence between the variables. I will discuss students' use and performance of different representations (i.e. graphs) in this study.

Batanero *et al.* (1996) suggest three factors which influence the difficulty of making association judgments, namely table size, students' previous theories about the context of the problem and the lack of perception of inverse association. Some students mixed correct intuitive conceptions with incorrect ones and chose an incorrect or partially correct strategy. Batanero *et al.* (1996) emphasise the complexity of a topic which is simple in appearance and summarise three categories of misconception of association judgment. First of all, some students have "unidirectional conception" (1996, p.164), so they judge inverse association as independence. Second, some students have a "deterministic conception" (1996, p.165) of association that assumes correlated variables should be linked by mathematical functions, and they do not admit exceptions to the relationship between the variables. Third, students who possess the "localist conception" (1996, p.165) compare the relative frequencies in only part of the data in the contingency table. Batanero *et al.* (2004) suggest that the meaning of any mathematical object consists of five interrelated elements. These are listed below, with a few adaptations of my own:

Problems and situations from which the object emerges;

Representations of data and concepts: including symbols, words and static or dynamic graphs which can be used to represent or manipulate the concepts and data;

Procedures and strategies to solve the problem;

Definitions and properties: such as meanings of parameters and probabilities within one, two or three standard deviations; and

Arguments and proof: such as informal arguments made on the basis of graphs (2004, pp.259-260).

Godino and Batanero (1998) argue that an understanding of a concept comes from an individual's meaningful practice with repeated solution of problems about the concept.

> The concept… has emerged and evolved progressively over time and practices created to solve problems. It has also generated some related concepts… Moreover, it has been the basis for developing new problem fields and tools for solving them (1998, p.184).

For instance, understanding normal distribution requires interrelated prior knowledge such as random variables, distribution, centre and spread (Artigue *et al.*, 2007). The application of the table of normal distribution involves interpreting graphical and symbolic representations. DelMas *et al.* (1999) find that students' misconceptions can be improved through conceptual change, and they therefore argue for testing students' predictions to confront misconceptions. I will follow this suggestion when I design the exercise questions.

2.1.2 Teaching and learning of statistical inference

Statistical inference is normally taught in the final part of the statistics curriculum, because it requires prior knowledge such as sample, population, mean, variation, and distribution, etc. Garfield and Ben-Zvi (2008) define statistical inference as “the theory, methods, and practice of forming judgments about the parameters of a population, usually on the basis of random sampling” (p.263). Statistical inference consists of parameter point/interval estimation and hypothesis testing, and involves two kinds of inference questions:

> generalizations (from surveys)… [which] is concerned with generalizing from a small sample to a larger population… and comparison and determination of cause (from randomized comparative elements)… [which] has to do with determining if a pattern in the data can be attributed to a real effect (Garfield and Ben-Zvi, 2008, p.263).

Its purpose is to answer particular questions, posed before the data are collected, and then make a formal conclusion.

Drawing inferences from data has become an important part of everyday life, but traditional approaches to teaching statistical inference (probability theory-based explanations in a formal statistical language) cause difficulties for students’ reasoning and understanding, so it is better to introduce the idea of inference informally at the beginning of the course by simulation or re-sampling methods to avoid over-mathematising (Garfield and Ben-Zvi, 2008). Some researchers, such as

Phannkuch (2006), Pratt (2008) and Zieffler *et al*. (2008b), investigate informal inferential reasoning (IIR).

There are many studies on students' difficulties of reasoning and understanding inferential ideas and procedures (such as students' common misconceptions about P-values and confidence intervals), but much fewer empirical studies on teaching concepts of inference (Garfield and Ben-Zvi, 2008). Liu (2005) finds that not only students but also teachers have difficulties in understanding statistical inference. Their difficulties consist of a variety of components. Thus,

> …students' statistical reasoning is often inconsistent from item to item or topic to topic, depending on the context of the problem and students' experience with the context (Garfield and Ben-Zvi, 2008, p.25).

In this study, I will examine whether the participants use consistent approaches to solve the same kinds of questions and whether they distinguish or get confused by the similar but different types of questions.

Some researchers offer suggestions for statistical inference teaching and learning. For instance, Watson (2004) shows that improving the instruction of sampling helps students better understand statistical inference. Rossman and Chance (1999) recommend ten principles for teaching the reasoning of statistical inference, such as always examining visual displays of the data. Garfield and Ben-Zvi (2008) argue that understanding how samples vary may help in making reasoned estimates and decisions, and that computation and formulas should be avoided to help students focus on the ideas of null and alternative hypotheses, P-value and types of errors. They suggest two pedagogical methods: modeling by the teaching of statistical reasoning and thinking in statistical inference (including making students' thinking visible from claims to conclusions) and "using argumentation metaphor for hypothesis testing", and they argue that students should be able to "use and develop their own statistical thinking,… rather than just solving a series of textbook problems for each given procedure" (2008, p.288).

Tarr and Shaughnessy (2007) find that some students learn formulas without understanding the meanings, and they perform badly when interpreting or applying information in tables and graphs. Chance *et al.* (2004) also find that some students can only develop a mechanical knowledge of statistical inference. This means that students can compute confidence intervals and carry out tests of significance, but they cannot necessarily explain related concepts, such as what *p*-value is. This study will further investigate how students solve questions in statistical inference with the tables of distributions and perhaps with graphs.

2.1.3 Theoretical frameworks for investigating students' understanding in statistics

Some theoretical frameworks have been created and applied in the research of statistics education. For instance, Kahneman's *et al.* (1982) heuristics, which includes representativeness and availability, explains students' misconceptions in probability. Researchers of statistics and probability have paid much attention to students' misconceptions, and recently some of them have changed their focus to the details of the development of students' understanding and thinking of statistics. Shaughnessy (2007) explains that

> ..most researchers in statistics… now hunt for a spectrum of student thinking in their research work, rather than taking an approach that students either do or do not "have it". This alteration… is partly due to an increased attention to and acceptance of a constructivist epistemology that places students at the active center of their learning (Shaughnessy, 2007, p.961).

With this shift of research direction, some frameworks and models for research in statistics were created to analyse statistics education from the three major perspectives of statistical literacy, statistical reasoning, and statistical thinking.

Models of statistical literacy help to identify what learners of statistics should do in order to succeed in this subject. For example, Curcio (1981) defines graph comprehension in three task levels: “a literal reading of the data… to read “between the data”… [and] to read “beyond the data”” (1981, p.22). Shaughnessy (2007) adds a fourth level, reading behind the data in order to introduce the context issues behind the graphs.

Chance *et al.* (2004) suggest a model of five developmental levels from merely knowing words and symbols without understanding them to completely understanding the process in order to collect detailed information of students’ statistical reasoning about sampling distribution. This model includes “Idiosyncratic Reasoning”, “Verbal Reasoning”, “Transitional Reasoning”, “Procedural Reasoning” and “Integrated Process Reasoning” (2004, pp.302-303).

Kirsch *et al.* (1998) suggest a framework of readers’ strategies to match information in reading a question. They use three kinds of process variables, namely type of information, type of match, and plausibility of distractors, and they also refer to Fry’s (1977) readability to measure the degree of difficulty of prose or document questions in large scale surveys. They stress that the given information and requested information must be first identified. In their exemplar questions, the questions (requested information) are independent sentences whereas the answers to the questions can be found explicitly or implicitly somewhere in the given texts. Once they understand what is requested, the respondents have to “[match] information in a question or directive to corresponding information in a text" (1998, p.114). This involves four strategies:

Locating: to find information according to the condition in the question.

Cycling: not only to locate and match some features, but also to engage in a series of feature matches to satisfy conditions of the question.

Integrating: to compare or contrast two or more pieces of information in the text when the information is provided in multiple sections.

Generating: to write a response by processing information from the text (1998, p 114).

When processing the text and making inferences, the respondent also needs to be aware of and identify the distractors (i.e. distracting information). Kirsch *et al.* (1998) use a coding system of scale to evaluate the task from the easiest (represented by 1) to the most difficult (represented by 5, 20, and 5 respectively in the three process variables). They suggest that this framework can be used to evaluate the degree of difficulty of document (including tables, signs, and graphs) and information organised in matrixes (i.e. columns and rows), with adequate procedural knowledge to "transfer information from one source… to another" (1998, p.119).

Kirsch *et al.* (1998) continue that requested information in quantitative questions is always an amount, so quantitative tasks involve not only readability, type of match and plausibility of distracters, but also two formulae variables: "operation specificity" (1998, p.122), which means the identification and choosing of the appropriate operation, and "type of calculation" (1998, p.123), which includes the type of arithmetic operation required and whether the operation should be performed alone or in combination.

Kirsch *et al.* (1998) developed the framework in order to evaluate the degree of difficulty of questions, but I believe that their process variables provide ideas to deconstruct the symbolic representation of the tables of normal distribution and t distribution. For example, the given information and requested information of the typical questions of using the tables of distribution such as $P(Z > a) = 0.025$ or $P(Z < -2) = b$ were contained in one equation, and I observed some of my previous students did not correctly distinguish the type of the given/requested information. The students dealt with the question according to what they interpreted from the question rather than what was actually given and requested in the question. Therefore, I decided to take advantage of Kirsch *et al.*'s framework to look into the steps of interpreting and solving the questions of using the tables of distribution and explore the causes of students' problems. I will explain how I used the framework to construct the study in Section 5.2.

Models of statistical reasoning help to explain how people think about statistics, what they know and understand and what difficulties they may have. Most researchers of statistics/statistics education have been concerned with students' reasoning on particular statistics concepts or processes by conceptual analysis. They attempt to "distill and categorize the ways the students reason with and think about the concepts" (Shaughnessy, 2007, p.966). For instance, Biggs (1979) suggests the five-level SOLO (Structure of Observed Learning Outcomes) taxonomy to code and analyse students' responses to particular tasks in clinical interviews.

Models of statistical thinking help researchers understand what concepts and processes statisticians feel are important. Some statisticians see statistical thinking as a "specific type of thinking that recognizes the variation around us and includes a series of interconnected processes" (Artigue *et al.*, 2007, p. 1042), and they therefore develop frameworks to describe it. For instance, Wild and Pfannkuch (1999) create a four-dimensional framework to organise some elements of statistical thinking during data-based enquiry. The four dimensions are described as follows:

Investigative cycle PPDAC (problem, planning, data, analysis and conclusion). This dimension deals with how an individual acts and what he/she thinks in a statistical investigation;

Types of thinking: including strategic thinking for all problem-solving and fundamental types of statistical thinking;

Interrogative cycle: a generic recursive process in all statistical problem-solving. The result of this dimension is like a distilling and encapsulating of ideas and information; and

A series of dispositions: which are personal qualities, such as curiosity and awareness, imagination and skepticism, etc. (1999, pp.225-235).

Wild and Pfannkuch (1999) explain that this framework can be used to analyse the thinking of not only statisticians but also students, and that anyone who is involved in enquiry can operate in all dimensions at the same time. Their arguments inspired subsequent research in assessing statistics reasoning and its components as a whole and the usage of transnumeration which is defined as "numeracy transformations made to facilitate understanding… [it] is a *dynamic* process of changing representations to engender understanding" (Wild and Pfannkuch, 1999, p.227).

In this section, I discuss several theoretical frameworks for investigating students' understanding in statistics. However, these frameworks are not designed to analyse solutions step by step and may not help in answering the research questions of the study. Therefore, I will have to develop new frameworks to analyse my data. The framework for the study will be introduced in Chapter 5.

Moreover, applications of graphs are also studied since distributions are usually displayed in graphical forms. I will discuss the literature of using graphs in the second part of this chapter.

2.1.4 Eight principles for learning statistics

In considering the three levels of understanding in learning statistics, Garfield and Ben-Zvi (2008) suggest eight principles for learning statistics. I will explain these principles and connect them to my study.

First, students learn by constructing knowledge: teaching is not simply telling, and learning is not simply remembering. That is, individuals will understand only after they construct their own meaning of the material. They interpret the new information according to their prior knowledge and connect what they learn to construct their own (new) meaning. However, the new meaning construction normally happens when the old meanings do not work or do not work well. Wild (2006) explains that statistics understanding depends on accumulation and inter-reaction of three elements, “context reality”, “statistical models” and statistical “knowledge” (2006, p.21). Using the tables of normal distribution (see Appendix 2.1. for Table A: left-tailed with negative and positive z scores and Appendix 2.2. for Table B: right-tailed with positive z scores) to find probability according to given z scores, and using the table of t distribution (see Appendix 2.3.) to find t score according to given probability were taught in the participants’ class as a basic skill. However, the teacher did not stress how to use the tables of normal distribution to find z score(s) according to a given probability in order to decide on the confidence interval or the critical region in statistical inference. Therefore, I will examine whether and how the participants developed such an ability and explore the factors leading to their difficulties.

Second, students learn by active involvement in learning activities: students learn better when they actively share ideas orally and in writing to and cooperating with other classmates to solve problems. Therefore, I will pay attention to the participants' interaction in the interviews and note any influences they seem to exert on one another.

Third, students learn to do well only what they practise doing: students learn better when they practise in new situations rather than simply repeating and reviewing tasks. In the interviews, the participants were first given typical exercise questions in different formats which they were supposed to be familiar with to see if they really understood, and they were then given questions in new formats to see whether their practice with the familiar questions helped them solve the new questions.

Fourth, it is easy to underestimate the difficulty students have in understanding basic concepts of probability and statistics. For example, the tables of distributions are taught and used in most statistics courses, but there is no relevant research about how students use the tables and what difficulties they may have. My study will focus on this issue.

Fifth, it is easy to overestimate how well their students understand basic concepts. This principle and the fourth principle may be the reasons why there is no research about how students use the tables of distributions and what difficulties they may have.

Sixth, learning is enhanced by having students become aware of their errors in reasoning. Students may be able to solve some formats of questions but not all of them. Giving students some questions in new formats which they have never encountered before and letting them compare different formats is likely to enable the researcher and students to find hidden misconceptions.

Seventh, technological tools should be used to help students visualise and explore data: they help students learn basic statistical concepts by providing different ways to represent the same data set. Although technical tools are not used in my study, the graphs of distribution can be drawn by hand to make a different representation. I will examine whether and how the participants used graphs to solve questions.

Eighth, students learn better if they receive consistent and helpful feedback on their performance. Therefore, I will discuss and explain to the participants their responses and mistakes in the exercise questions before the interviews finish. By doing this, I hope students in the study can realise any problems they have and improve their performance. It will also enable me to observe whether the participants make the same mistakes in the subsequent interviews.

2.1.5 Challenge of statistics education

Shaughnessy (2007) argues that the focus of research in statistics has moved from probability concepts and probability distributions to students' understanding of specific statistics concepts, and

> [p]erhaps the overarching goal of statistics education is to enable students… to read, analyze, critique, and make inferences from distributions of data (Shaughnessy, 2007, p.968).

Since distribution involves several pre-requisite statistical ideas such as centre (average), spread (variability) and shape, many studies are devoted to investigating students' reasoning in these aspects. For example, Garfield and Ben-Zvi (2008) discuss how students learn to reason about distribution, centre, variability, comparing groups and statistical inference.

Some ideas in the statistics curriculum regarded as basic are actually complicated, and errors and misconceptions can also occur when dealing with simple ideas. For example, Lecoutre (1998) finds that some suggestions to complement a statistical test with confidence interval were based on misunderstandings, and that “the “correct” frequentist interpretation of confidence intervals… does not make sense for most of the users” (1998, p.243).

Reid and Petocz (2002) notice that students who take several statistics courses do not necessarily understand the purpose of statistics at a higher level. Zieffler *et al.* (2008a) find that many students cannot correctly reason on important statistical ideas, and their misuses and misconceptions still exist even after formal instruction with good methods. They also find that traditional effort-based learning approaches such as homework assignments may result in lower correct reasoning abilities, and they thus suggest that teachers bring interesting examples and encourage students to feel the benefits learning statistics. When students’ attitudes are improved, they may be more willing to attend and enjoy the class, and their learning will improve as a result. I will explore each participant’s misuses and misconceptions of each exercise question, explain these at the end of the clinical interviews, and test the ideas again in the subsequent interviews to see if the participants still have the same problems. Also, I will consider whether and how the participants’ attitudes to visualisation influence their performance on the exercise questions. Applications of graphs are studied since distributions are usually displayed in graphical forms. I will discuss the literature of using graphs in the second part of this chapter.

2.2 Visual thinking

In my study of solving questions of statistical inference with tables of distribution, I felt it necessary to consider the issues of visualisation because some properties of an image, such as symmetry, skewness, convergence or divergence, can be displayed and observed in the visualisation directly (Hoffman, 1987). On the basis of Zimmerman and Cunningham's (1991) and Hershkowitz *et al.*'s (1989) arguments, Arcavi (2003) defines visualisation as

> the ability, the process and the product of creation, interpretation, use of and reflection upon pictures, images, diagrams, in our minds, on paper or with technological tools, with the purpose of depicting and communicating information, thinking about and developing previously unknown ideas and advancing understandings (p.217)

Brown and Liebling (2005) believe that successful learners take advantage of visualisation. They suggest that those working to solve problems should spend some time to imagine and visualize, as this may help them to check whether the problem is really what they had thought it was. Corter and Zahner (2007) suggest that visualisation, visual solution methods, and experimenter-provided external visual representations are helpful in solving problems. Manipulation of visualisation has proved beneficial in problem-solving in some aspects, such as the role it plays in the processing of information (Lowrie, 2001; Dreyfus, 1991). In addition, it "give[s] us a very helpful mental support for the key ideas" (Guzmán, 2002, p.6) of inverse image of a mapping, it "evokes the idea… plays a very important role as evocative element of knowledge" (Guzmán, 2002, p.8), and it makes it possible to avoid unnecessary calculi or to control and predict in algebraic work (González-Martín and Camacho, 2004).

Corter and Zahner (2007) explain that “using multiple representations of a problem (including visual ones) leads to a fuller understanding of the problem and an increased “depth of processing” (p.24) because the act of restructuring is an important step in dealing with novel problems. Schwartz and Martin (2004) also find that experimenter-prompted graphical invention activities obviously help students realise statistical concepts.

However, Corter and Zahner (2007) point out that each user of a graph may benefit to different degrees. For example, spatial visualisers perform better than others. Furthermore, using graphs cannot guarantee successful problem-solving, and there may be a negative correlation between the proportion of graph use and the number of correct answers because some students attempt to use a graph only when they feel they are addressing difficult problems. Moreover, students do not always find the correct graphical representations, and a wrong or inappropriate representation may cause an incorrect answer. Duval (1999) points out that “learning visualization in mathematics is not quite as easy and successful as it is for physical objects and real environment” (p.15). Although people tend to say that an image is worth a thousand words, it has to be based on an assumption that images are “correctly deciphered and understood” (Guzmán, 2002, p.4); otherwise the images are worth nothing, and incorrect visualisation only leads users to errors.

In the second part of this chapter, I will briefly introduce a map of the development of visualisation in learning and teaching mathematics and studies on different forms of visualisation such as mental images, graphs and tables. I will also introduce graph comprehension and four critical factors which influence graph comprehension. Later, I will discuss graph sense and its basis (number sense and symbol sense). In the last part of this section, I will talk about the application and reluctance to use visualisation.

2.2.1 Development of visualisation in learning and teaching mathematics

Traditional mathematical research focuses on bodies of mathematical theories, but in recent decades some researchers have shifted their attention to mathematical practice. Giaquinto (2005) suggests four mathematical activities (with goals), namely: discovery (for knowledge), explanation (for understanding), justification (for relative certainty) and application (for practical benefit). He explains that each activity has three sub-activities: making (mainly by mathematicians), presenting (by mathematicians and teachers), and taking in (by mathematicians, teachers and students). Later formulation is added to the list, and this leads to the "invention of symbol systems and associated algorithms for problem solving" (2005, p.85). Also, he constructs representation including lingual symbol systems and visual diagram systems in the list of mathematical activities.

> Language is just one of the many forms in which information can be couched. Visual images… are other forms (Mancosu, 2005, p.24)

Giaquinto (1994) argues for multiple roles of visualisation. For instance, an individual may visualise to strengthen his/her grip on something independently known. He stresses that "in analysis visualizing may be heuristic useful, but it is not a means of discovery" (1994, p.790), and that "visual thinking in analysis is often an aid to understanding" (1994, p.811). Mancosu (2005), on the other hand, argues that visual representation not only enables discovery without proof but also helps explanations which aim to make something already known more intuitive. He explains that

> …epistemic function of visualization in mathematics can go beyond the merely heuristic one and be in fact a means of discovery… one discovers a truth by coming to believe it independently in an epistemically acceptable way… [probably] through

the process of visualization (Mancosu, 2005, p.22).

is a long tradition of research on “mental imagery in all of the sense modalities” (Presmeg, 2006, p.205) of sight, hearing, smell, taste and touch and their interactions in psychology, but visualisation has long been treated as “certainly *not*… significant” (2006, p.205) in the area of mathematics education. However, when the constructivist and qualitative research methodologies blossomed in the 1980s, visual thinking and qualitative research became

> a suitable vehicle for investigating… thought processes associated with the use of mental imagery and associated forms of expression in learning mathematics… [because] mathematics is a subject that has diagrams, tables, spatial arrangements of signifiers such as symbols, and other inscriptions as essential components (Presmeg, 2006, p.206).

Eventually, the value of imagery and visualisation was noticed, and visualisation was listed as a separate research topic for the first time in PME-15 in 1991.

Presmeg (2006) adds pattern imagery to Dörfler’s (1991) four visualisation-relevant “image schemata with their “concrete carriers”” (2006, p.208), namely figurative, operative, relational and symbolic image schemata. She explains that pattern imagery is a “strong source of generalization” (2006, p.209) and that it is capable of depicting relations. In her study, Presmeg (2006) finds that mathematical difficulties always relate “in one way or another to problems with generalization” (2006, p.209), so she suggests using pattern imagery and metaphor via an image to help the static image become the bearer of generalised mathematical information for a visualiser. However, she points out that

> …concrete imagery needs to be coupled with rigorous analytical thought processes to be effectively used in mathematics (Presmeg, 2006, p.209).

Although some teachers in Presmeg's (1991) study believe that the visual mode is not really necessary and can be dispensed with after reaching the goal of generalisation, Presmeg (2006) insists on the importance of "personal metaphors, encapsulated in imagery, not only for individual meaning making and retention, but also in the service of mathematical generalization" (2006, p.214). Giaquinto (1994) suggests that visualisation helps us not only to "locate roughly values for given inputs", but also to "avoid mistakes" (1994, p.811). However, knowledge of the properties of diagrams will not necessarily grow with a learner's age (Diezmann, 2005), so the quality of mathematical visualisation should be improved by education.

Mancosu (2005) distinguishes between two visualisation conceptions, namely "mental imagery" (visualisation by means of mental images) and "diagrams" (visualisation by means of drawn or computer-generated images) (2005, pp.13-14). Presmeg (2006) defines a visual image as a "mental construct depicting visual or spatial information" (2006, p.207) (e.g. geometrical diagrams, maps, graphs or visual scenes of real-world situations) and a visualiser as a "person who prefers to use visual methods when there is a choice" (2006, p.207). She explains that visual images may appear in a person's mind and guide in creating a spatial arrangement. Therefore,

> visualization… include[s] processes of constructing and transforming both visual mental imagery and all of the inscriptions of a spatial nature… Inscriptions [i.e. graphical representation] are central to scientific practice (2006, p.206).

2.2.2 Graphs and tables relevant to the study

Statistics and data analysis became important in school curricula during the 1990s, and data analysis largely relies on graphical representations, especially when creating graphs has become easy with the technology available in schools. Therefore, the ability to read, understand, analyse and interpret information presented in graphs or tables becomes important.

> Graphs are critical for data representation, data reduction and data analysis in statistical thinking and reasoning (Shaughnessy, 2007, p. 988).

Friel *et al.* (2001) combine several aspects of studies on graphical representation and explain that business graphics can be an aid in decisions, that graphs involve audiovisual communications, and that graphs involve human interaction with a display to propose a framework to discuss graph comprehension. Graham (1987) recommends Statistical Investigation in the Secondary School (SISS) as a pre-statistics course. The statistical investigation includes four key stages, namely posing the question, collecting the data, analysing the data and interpreting the results. The SISS course consists of simple graphical methods for displaying, representing and interpreting data, and all of this is done without the use of formulae, which helps to encourage common-sense interpretation of data and develop intuition of statistical thinking. It also avoids simply substituting numbers into meaningless formulae. Useful graphical methods for statistical investigation include stemplot, histogram, scatterplot, bar chart, pie chart and contingency table. For instance, the ideas of distribution and distribution curve are generated from the ideas of average and variation, and histogram.

A graph is "information transmitted by position of point, line or area on a two-dimensional surface" (Fry, 1984, cited in Friel *et al*. 2001, p.126). Kosslyn (1994) suggests that graphs consist of several structural components, such as a framework (axes, scales, grids, reference markings) which gives information about the used measurements and measured data; labels naming the measurement and measured data; contents such as specifiers representing data values; and the background of a graph such as coloring, a grid and pictures which can be superimposed on the graph.

Most statistical graphs are displayed using L-shaped coordinate systems, and quantities are represented by spatial characteristics such as height or length. According to Friel *et al.* (2001), most currently used statistical graphs including picture graphs, line plots, bar graphs, pie graphs and histograms were developed by Playfair in the late 1700s. Stem(-and-leaf) plots and box(-and-whisker) plots were introduced by Tukey (1977) to make schematic summaries.

Some particular statistical graphs, such as bar graphs, line graphs and pictographs, are found to be difficult for students to read and interpret. Integrating information in graphs or summarising a conclusion from graphs is also found to be difficult for students. Shaughnessy (2007) suggests proportional thinking as the key to making sense of statistical graphs, as it can help students shift from additive to distributional thinking. The tables of continuous distribution are taught and used to find the probability or value in particular situations, such as P(Z > -2), $z_{0.05}$, $t_{0.01}(10)$, $F_{0.05}(8,15)$, or $\chi^2_{0.05}(10)$. Each probability in the tables of distribution is equal to the representing area in the graph and involves the application of integration. Dunham and Osborne (1991) stress that

> …integration has intuitive definitions rooted in the graph of functions. A thorough understanding of how functions relate to their graphs is crucial to understanding the

concepts of calculus (1991, pp.36-37).

Dunham and Osborne (1991) notice that researchers, teachers or students who have successfully made connections between the representations feel that graphs are obvious and trivial, so they do not find difficulties or false notions about graphs. Thus, they suggest three types of behaviour to build better visual intuition to help students learn “how to see” (1991, p.37). These three behaviours are “making connections between symbolic and graphical representations”, “concern[ing] scale and scaling”, and “transformations of functions” (1991, p.37). Each of these behaviours is not independent of the other two, but rather affects performance in them. I will explain these behaviours as follows.

Dunham and Osborne (1991) explain that most problems in visualisation are caused by the “failure to make connections between symbolic and graphical representations of functions” (1991, p.37). They find that three fundamental understandings of the nature of graphs and functions are required to make connections. First, they stress “the notion of what an ordered pair is and how an order pair is incorporated into thinking about a function” (1991, p.37). An ordered pair within the context of a function carries not only the location of a point, but also the idea of projection from the domain to the range. Second, some students do not see a graph as a visual representation of a mapping from the domain to the range, just as they see a function as a symbolic representation of the same mapping. Third, students often “focus on isolated, distinct points of a graph… and fail to draw inferences from visual characteristics that are representative of interval behavior” (1991, p. 38). That is, they disassociate the numerical characteristics of domain and range from the visual aspects.

It has also been suggested by Dunham and Osborne (1991) that one of students’ major visualisation difficulties involved in graphing is their projection from the critical coordinates of the multiple points onto the x-axis. Thus, they recommend skills such as placing a piece of paper on the computer screen with a corner on the point and the edges of the paper horizontal/vertical to the x- and y-axes, or drawing an arrow from a point to the axes to help students build ideas of projection. They argue that such approaches are still needed in university-level instruction, and one possible difficulty of students' ability to understand is that they do not understand the information carried by the symbol $f(x)$, and do not see it as a specific value. Therefore,

> [m]aking explicit the meaning of the separate symbol referents within the $f(x)$ notation improves understanding significantly to help students connect the visual and the symbolic (Dunham and Osborne 1991, p.39).

Another major difficulty found by Dunham and Osborne (1991) was that many students see a graph as "a string or a rubber band separate from the function it represents" (p.40). Therefore, they encourage students to see a function as a "collection of ordered pairs" (1991, p.40). This problem may be caused by the confusion of natural simplification of language. Students are expected to build the idea that a curve is "a collection of points rather than a single entity in and of itself" (1991, p.40), perhaps simply by using pencil-and-paper techniques. In addition, students cannot "infer characteristics of a function from more than one point of a function" (1991, p.41). However,

> [t]he failure to connect symbolic and graphical representations is correctable if instructors anticipate difficulties of students. Several of the corrective moves… are language driven. .… If a student understands contextual cues well enough to use the words appropriately in both [symbolic and visual] representations, the student can visualize and learn from graphs more readily (1991, p.41).

The second type of behaviour concerns scale and scaling because “scale is critical to successful use of graphing technology” (Dunham and Osborne, 1991, p.42). However, scale was not regularly noticed by students. Such inattention to scale makes it difficult for students to “[interpret] asymmetric scales and [choose] appropriate scales to make good use of the graphing space” (1991, p.43). Each probability value in the tables of distributions can be represented by and is equal to the shadow area in the graphs. However, there is no scale or grid for the graph users to measure the shadow areaWhether the students remember the principle that the total area below the curve equals one and represent the area correctly is worth investigating, and this study will address this issue.

The third behaviour is associated with transformation of function and change of scale. Very few students notice the necessity of scale when drawing graphs. Even if they notice or are told about the change of scales, they still see “the figure changing independently of the coordinate system” (Dunham and Osborne, 1991, p.45) because they mix the ideas of geometric transformation with scale changes. Such confusion is complicated by the fact that the “shape of the graph depends on the viewing rectangle under consideration” (1991, p. 45). This situation can be improved by training students “to recognize critical visual features of a graph, to relate these to the global setting of the graph, and to track the features under transformations and scale changes” (1991, p.46). Dunham and Osborne (1991) suggest giving students an opportunity to discuss and listening to them talk about graphs in order to understand what and how students think. They add that instructors must teach students “*how* to see what a graph reveals (1991, p.46)”, rather than paying attention solely to symbolic manipulations.

However, the reading and interpreting of the table of normal distribution and the normal curve have not been discussed in previous research. Friel *et al.* (2001) point out that "the structure of tables can be linked to the structure of graphs" (2001, p.127). They explain that tables are used to display data or organize information as transition tools to create graphical representations. Some graphs are related to other graphs, and some can be developed into other graphs. The transitions among graphs are found to be a benefit to understanding (Bright and Friel, 1996). In fact, the information in symbolic form can be expressed by the graph of normal distribution, and the transferring principles which students largely relied on to solve questions of statistical inference can also be explained by the properties of graphs, such as direction, location and area (transferring principles will be discussed in Chapter 5).

2.2.3 Graph comprehension and four critical factors

Curcio (1987) suggests three abilities of a graph reader, namely literal reading of data in graphs, interpreting, and generalising from the data. Friel *et al.* (2001) define graph comprehension as a "graph reader's abilities to derive meaning from graphs created by others or by themselves" (2001, p.132). They argue that graph comprehension includes not only "reading and interpreting graphs", but also "graph construction or invention or graph choice" (2001, p.129). Kosslyn (1985) argues that comprehension involves not only "simply recognizing parts and relationships among them in a display", but also the reader's ability to "draw inferences about the actual information content" (1985, p.508). Jolliffe (1991) suggests three behaviors of comprehension of information: translation, interpretation and extrapolation.

> Translation is an activity requiring the change of form of a communication, such as describing the contents of a table of data in words or vice versa, or being able to interpret a diagram at a descriptive level… Interpretation involves a re-arrangement of material. Extrapolation is an extension of Interpretation to include statements about the consequences of a communication. (Jolloffe, 1991, p.462).

Friel *et al*. (2001) suggest that looking for relationships between specifiers in a graph or between a specifier and a labeled axis can help to interpret graphs and sort the important factors from the less important ones.

Curcio (1981) defines graph comprehension in three question levels: elementary level (reading the data), intermediate level (reading between the data) and advanced level (reading beyond the data). Reading beyond the data is more difficult because it "requires eliciting and evaluating opinions (rather than facts) about information presented in representations" (Friel *et al.*, 2001, p.132). Shaughnessy (2007) adds a fourth level (reading behind the data) to consider the context issues behind the graphs. Friel *et al.* (2001) focus on statistical graphs and suggest four critical factors which influence graph comprehension. These factors consist of the purpose for using graphs, characteristics of tasks, characteristics of the discipline and characteristics of readers.

Purpose for Using Graphs

Kosslyn (1985) argues that there are two types of visual displays; some of them are used for data analysis and others for communication, and both types can be used to store information. Graham (1987) observes the following:

> …it is difficult to choose a suitable representation for data if you don't know the purpose for which the representation is required. Choosing a graphical method… depends not only on *what* sort of data you have collected but also on *why* you have collected them (1987, p.10).

There are several ways to categorise data. For instance, Graham (1987) simplified Stevens' taxonomy of four scales of measures (naming scale, ordering scales, interval scales and ratio scales) to three scales: category, discrete variable and continuous variable. It is important to recognise and distinguish between the types of data because some representations are suitable for particular data type(s) but not suitable for other type(s). For example, stemplots are suitable for discrete and continuous variables but not for categories. Bar charts are used to display discrete data, while histograms are used to represent continuous data. Graham (1987) argues that "why you have collected your data will help you to choose an appropriate representation" (1987, p.31). He suggests four purposes for collecting data: describing, summarising, comparing and generalising. In fact, the graph of normal distribution, no matter if it is right-tailed or left-tailed, is not the only static one. The graph with varying z score and probability summarises hundreds of values listed in the table of normal distribution, and this enables users to compare values and generalise. However, many students do not use graphs, and many users do not vary the location of z score and area of probability. Consequently, they cannot benefit from the graphs. This also happens to graphs of other continuous distributions.

Friel *et al.* (2001) add that graphs used for data analysis can serve as "discovery tools" (2001, p.132) to make sense of data in the early stage of analysis. Similarly, Tukey (1977) argues that

> **[t]he greatest value of a picture** is when it *forces* us to notice **what we never expected to see** (1977, p. vi).

Friel *et al.* (2001) suggest that graph instruction "within a context of data analysis may promote a high level of graph comprehension" (2001, p.133). However, the graphs in textbooks are usually preprocessed, so some students may be able to graph but do not know why. I will consider how the graphs in the textbooks influence students' learning in this study.

For another purpose, Kosslyn (1985, 1994) suggests that “[g]raphs should be constructed so that one notices the more important things first” (p.504) and that “a good graph forces the reader to see the information the designer wanted to convey” (1994, p.271). Kosslyn (1985) explains that displays should be subject to only one interpretation and should not lead readers to inappropriate information. Many students use only one graph for many questions in different conditions, and the wrong graphical representation does not help them understand the situation. This study will also be concerned with this issue.

The graphs of the normal curve represent important information in different aspects. For instance, z score is represented by its location on the x-axis (not merely the length from zero to it), relation is represented by the direction of the shadow area, and probability is represented by the size of the shadow area. Kosslyn (1985) suggests a principle of graphical representation in which “[g]reater quantities are indicated by “more” of a mark (higher lines or bars, larger areas, etc.” (1985, p.509). Friel *et al.* (2001) define graphs used for communication as “pictures intended to convey information about numbers and relationships among numbers” (2001, p.133). Graphs used for communication are also relevant to the other three factors which influence graph comprehension.

Characteristics of Tasks

Legge *et al.* (1989) argue that graph perception involves visual perception in analysing graphs. Therefore, Friel *et al.* (2001) suggest three tasks in the perceptual process: visual decoding (visual design), judgment and context. Cleveland (1985) argues that a focus of displaying data is the accuracy of judgment. For example, reading scales in bar graphs is more accurate than judging angles in pie graphs, and judgment of position, length and angle is more accurate than judgment of area and volume.

Vessey (1991) points out that a graph and table may contain the same information, but graphs emphasise spatial information while tables emphasise symbolic information. Therefore, tasks can be divided between spatial tasks (using graphs) and symbolic tasks (using tables), and students perform well when "there is a cognitive fit (match) between the information emphasized in the representation type and that required by the task type" (1991, p.219)—that is, when graphs support spatial tasks or when tables support symbolic tasks. Graham (1987) adds that the use of statistical graphs should be decided by one's purpose for data analysis, such as describing or summarising, comparing or contrasting data sets, and generalising populations or predicting, and that each kind of graph has its own usage and constraint. Simkin and Hastie (1987) argue that the interaction of judgment tasks and visual decoding influences graph-comprehension performance. Also, improper graphs cannot help users to realise the situation.

Characteristics of the Discipline

Moore (1991) argues that statistics is about data which are numbers, but not just numbers. He explains that

> *Data are numbers with a context*… Statistics uses data to gain insight and to draw conclusions… [with the tool of] graphs and calculations… guided by ways of thinking that amount to educated common sense (p.xxv).

Therefore, doing statistics is more than simply manipulating numbers, and statistics learning usually involves collecting, describing and presenting data and then drawing conclusions from the data (Moore, 1991). Friel *et al.* (2001) suggest that graph comprehension may be influenced by spread and variation of data sets, type and size of data sets, and complexity of the representation.

Spread and variation. L-shaped statistical graphs are often used in data reduction in order to structure information. Axes may have different meanings, such as the values in vertical axes of each observation of ungrouped data or the frequencies in vertical axes of each observation or group. The vertical axes (y-) usually represent values or frequencies, whereas horizontal axes (x-) display observations or groups. Dunham and Osborne (1991) stress that scale and scaling are important in interpreting data, such as that graph scale may influence one's reading of the frequency of values. Rangecroft (1994) points out that choosing a proper scale for data is usually more difficult than drawing or reading given scales. It also seems to be difficult for students to locate a given z score or decide the size of the shadow area of probability. This study will address this issue by examining how students decide the scale and location on the x-axis and the area below the curve of distribution.

Data type and size of data set. Type and size of data set should be considered when choosing graphs to represent the data. For instance, bar graphs can summarise data and pie charts can be used to compare percentage.

Graph complexity. Teachers should help students to build a fundamental notion of comparing frequencies or measures from objects to abstract graphs, and then to develop scaling techniques. Some graphs, such as histograms, box plots and line graphs, are more difficult than others.

Characteristics of Graph Readers

Berg and Philips (1994) find that graphing ability has a positive relationship with logical thinking and proportional reasoning. Students with proportional reasoning perform well in graphical situations, whereas students without logical thinking are less capable of interpreting or constructing graphs. Carpenter and Shah (1998) explain that individual graphic knowledge also plays as important role in the comprehension process.

Roth and McGinn (1997) stress meaningful practice and experience of graphs; they suggest that when students simply make graphs that lack meaning, this does not help them learn graphs well. Furthermore, they explain that lack of competence may be caused by lack of experience and participation, rather than exclusively cognitive ability. Curcio (1987) argues that mathematics knowledge and experience are both important to graph comprehension, and that the mathematical contents of a graph, including number concepts, relationships and fundamental operations, are necessary prior knowledge for graph comprehension.

Roth and McGinn (1997) argue that students should collect real-world data to construct their own graphs and verbalise the relationship between the data, such as “larger than” or “twice as big as”, to enhance their mathematical knowledge. Gal (1993) finds that applying basic proportional concepts to numerical data presented in statistical contexts is difficult for some students. It may be helpful to let students collect data, build a histogram and then transfer the histogram to the normal curve. This may enable them to seriously consider the meaning of positive/negative z score and bigger than/smaller than when learning statistical inference.

2.2.4 Number sense, symbol sense and graph sense

Graph comprehension involves not only reading and making sense of already constructed graphs in daily life, but also understanding what is required in graph construction and what the optimal choice is. It is one “way of thinking rather than as bodies that can be transmitted to others” (Friel *et al.,* 2001, p.145). For instance, Friel *et al*. (2001) suggest that graph sense is the gradual result when one creates graphs or uses already designed graphs in a variety of problem contexts that require making sense of data. They explain that

> perceptual demands related to graph design affect graph comprehension (Friel *et al.*, 2001, p. 149).

Graph sense was developed on the basis of number sense and symbol sense. These three senses are involved in this study, so here I will discuss how they relate to the study.

Number sense

Anghileri (2000) explains number sense as a "'feel' for numbers" (2000, p.1) by giving some examples: children can become aware of how each number relates to others according to their experience, such as six is after five and before seven (as their age), two and two and two (result of throwing three dice), four and two (counting wheels of a car and a bike) or double three (another result of throwing two dice). Also, if an individual with number sense calculates the result of adding 'one and two and three and… and ten is fifty-five', he/she can get the result of 'one and two and three and… and ten and eleven' as 'fifty-five and eleven', rather than by adding all of the numbers again. Number sense is "highly personalized" (2000, p.3) and involves not only what ideas numbers are about, but also how they are built. Anghileri (2000) argues that number sense is people's ability

> to make generalizations about the patterns and processes they have met and to link new information to their existing knowledge… [It is] an awareness of relationships that enable the subject to interpret new problems in terms of results they remember (Anghileri, 2000, p.1).

Friel *et al.* (2001) suggest that students represent numbers initially by using concrete physical objects, then by using pictures or materials, and eventually by using more abstract representations such as bar graphs or line graphs. Tables can also be used together with graphs for data representation or data organisation. The present study involved two tables of normal distribution and the table of t distribution. There were some patterns between the numbers in the tables. For instance, Table A includes negative z scores and positive z scores, and the numbers were not randomly listed in it. The z scores were presented in the table by a particular way of combination, such as -2.35 was presented as -2.3 in the left column and .05 on top. The z scores gradually increased from the top to the bottom, and the negative z scores increased from the right-hand side to the left-hand side, whereas the positive z scores increased from the left-hand side to the right-hand side. The least z score was -3.99 in the upper-right corner of the first half of Table A, and the largest z score was 3.99 in the bottom-right corner of the second half of Table A. On the other hand, the probabilities also increased with the z scores, but there was no negative probability, and the probability was from 0.0000 to 1.0000.

The numbers of Table B also had some tendencies, but Table B only had positive z scores. The z scores also increased from the upper-left corner 0.00 to the bottom-right corner 3.99, but the probabilities decreased in the direction from 0.5000 to 0.0000.

The table of t distribution including two pages was in a different format, presenting v in the left column and probabilities on top. The probabilities decreased from .25 in the left-end to .00005 in the right-end, and the v increased from 1 on top to 120 on bottom. The corresponding t scores also decreased from the upper-right corner to the bottom-left corner.

Being aware of the tendencies of the values in the tables, the user should be able to find a value by comparing it to the value he/she was looking for and the values he/she has found. For example, if he was looking for z score -2.35 in Table A and he was looking at -2.5, he should look at the lower part of the current location, rather than looking for it everywhere.

Negative numbers have been found to be more difficult than positive numbers for students. How negative numbers should be introduced to students is still a question for educators. Some educators prefer formal abstraction, while a larger number of educators prefer to introduce the idea through models and/or concrete representations. Negative numbers are frequently extended from natural numbers to integers, or from positive real numbers to negative real numbers. However, students perform better on questions involving negative numbers in a story than on similar questions that are purely symbolic (Verschaffel *et al.*, 2006). I will examine how students perform on symbolic and narrative questions involving negative numbers.

Symbol sense

According to Fey (1990), symbol sense is a "comparable informal skill required to deal effectively with symbolic expressions and algebraic operations" (1990, p.80). An individual with symbol sense should be able to perform the following (adapted and abbreviated from Fey):

Scan an algebraic expression to make rough estimates of the patterns that would emerge in numeric or graphic representation;

Compare orders of magnitude for functions with rules of the form $n, n^2, n^3, \ldots,$ and k^n (*sic*). This skill is a bridge between number sense and symbol sense;

Scan a table of function values or a graph, or interpret verbal statement, in order to identify the likely form of an algebraic expression;

Inspect algebraic operations and predict the result, or inspect the result to judge whether it has been performed correctly and

Determine which equivalent form is appropriate to answer a particular question (1990, pp.80-81).

These abilities were needed for using the tables and graphs of distribution. Taking the typical format representing z score and its corresponding probability for example, after an individual with symbol sense was taught $P(Z > 0.67) = a$ and $P(Z < b) = 0.67$, he/she was supposed to be able to distinguish between the two formats and interpret 'a' as a probability and 'b' as a z score. If he /she used a graph, he/she should know that 'a' equals the corresponding shadow area in the graph. In addition, he/she should be able to display the direction of '>' as right-tailed and '<' as left-tailed in the graphical representation. Moreover, he/she was expected to be able to mark the z score '0.67' on the positive (right-hand) side of the x-axis and display the probability '0.67' as a shadow area below the z curve. The shadow area of '0.67' should occupy more than half of the space below the z curve since the total area below the z curve is 1. I will take these issues into account and examine how the students handle this when I design the study.

Graph sense

The goal of a display is to make sense of information as easily as possible, and students should use their own ways of representation for either creating or making sense of data (Friel *et al.*, 2001). Students can construct and develop their own meaning or knowledge through discussing their strategies and making sense of others (Anghileri, 2000).

Shaughnessy (2007) finds that students interpret graphical representations badly. They usually cannot reason beyond graphs, and some of them cannot even read a graph at all. He suggests two methods to explore the components to develop graph sense. One method is to investigate students' graph comprehension and graph interpretation skills, and the other is to analyse the student-generated graphs. Therefore, I will collect and analyse the participants' actively drawn graphs and ask them to draw some graphs for comparison and analysis.

Friel *et al.* (2001) suggest six behaviours associated with graph sense. These six behaviours correspond to Curcio's (1981) three levels of graph reading, and they are also needed to connect the table of normal distribution with the graph of the normal curve:

1. To recognise the components of graphs, the interrelationships among these components, and the effect of these components on the presentation of information in graphs;
2. To speak the language of specific graphs when reasoning about information displayed in graphical form;
3. To understand the relationships among a table, a graph, and the data being analyzed;
4. To respond to different levels of questions associated with graph comprehension or, more generally, to interpret information displayed in graphs;
5. To recognise when one graph is more useful than another on the basis of the judgment tasks involved and the kind(s) of data being represented;
6. To be aware of one's relationship to the context of the graph, with the goal of interpretation to make sense of what is presented by the data in the graph and avoid personalization of the data (Friel *et al.*, 2001, p.146)

Therefore, after being taught the table of normal distribution and its graph, an individual with adequate graph sense is expected to understand that

There are two areas below the z curve. Both the shadow area and the other area will vary when the z score changes but the sum of the two areas remains 1;

The z score (z_α), the size of the shadow area (α) and the direction of the shadow area (left-tailed or right-tailed) can be represented by symbolic form $P(Z < z_\alpha) = \alpha$ or $P(Z > z_\alpha) = \alpha$;

The symbolic form, the graphical representation and the table of normal distribution are not isolated from but related to each other. In other words, each of the three formats can be represented in two other forms, and when one format changes, the other two formats also change;

The value on the z-axis is a z score, the areas below the z curve are probability values, and the direction of the shadow area indicates that the given area represents the probability of Z which is 'bigger than' or 'smaller than' the z score;

The graph of z distribution can be drawn according to each z score or probability. The direction of the shadow area can be decided according to the format of the table. Besides, the graph also can be drawn according to the symbolic form.

In terms of the issues raised above, I will examine whether the students used or drew graphs when using the tables of distribution, how they used or drew graphs, and in what aspects they have problems with the graphs.

2.2.5 Application and reluctance

There are two types of mathematical thinking: "tendency towards abstraction" and "tendency towards intuitive understanding which stresse[s] processes of visualisation and imagery” (Breen, 1997, p.97). Although traditional curricula in many countries rely on the tendency towards abstraction, Breen (1997) emphasises that the mathematical and educational imageries are important tools for learning. Wheatley and Brown (1994) argue that imagery plays an important role in mathematical reasoning. Gray (1999) explains that imagery includes not only what is essential to thought but also what generates thought.

Although visualisation has been proved helpful in mathematics learning by many researchers, it is still the case that very often mathematics students are used to symbolic representations and reluctant to use “visual approaches to support meaningful learning” (Healy and Hoyles, 1996, p.67), even if particular images or diagrams can help mathematical generalisation and link different modes of thinking. Presmeg and Bergsten (1995) find that students’ preference for mathematical visualisation follows normal distribution. A few of them, influenced by socio-cultural factors, always feel that visualisation is necessary, and some others never felt that visualisation is necessary.

Dreyfus (1991) explains that students are reluctant to use visualisation because of its low status in the mathematics classroom, so he suggests the need to lift the status of visualisation from a “helpful learning aid” to a “fully recognized tool for learning and proof” (p.33). The present study will examine the role and status of the graphs of distribution when learning to use the tables of distribution.

In the first part of this chapter, I introduced the development of research of statistics education in the past and present, some principles for learning statistics, theoretical frameworks for analysing data and challenges in statistics education. I also addressed the teaching and learning of statistical inference, which this study will focus on. In the second part, I introduced the development of visual thinking and how visualisation helps learning mathematical subjects. I also considered factors of graph comprehension and discussed how number sense, symbol sense and graph sense relate to this study. In the next chapter, I will introduce the educational systems in Taiwan along with the backgrounds of the site of data collection and the participants.

CHAPTER 3

CONTEXT

3.1 Educational systems in Taiwan

According to the Ministry of Education in Taiwan, six years of elementary school and three years of junior high school are compulsory for every child between the ages of 6 and 15 years. Children or parents who delay or avoid compulsory education can be prosecuted. Children delayed beyond the legislative age for compulsory education still have to finish the elementary and junior high school studies.

There are two main categories for education after compulsory education: higher education and technical and vocational education. Taking joint entrance examinations for each education system is the only way for junior high school graduates to enter the three years of high school, three years of vocational high school or five years of professional school. Taking entrance tests is the traditional way to enter a new educational institute after compulsory education, but application for admission to enter universities was introduced as an alternative method in the late 1990s, and this method is used by more and more students now.

Normally students in the two categories would follow the two standard routes. High school graduates may take university joint examinations to enter universities, and if they want to get a master or doctoral degree, they have to take an entrance examination again which is held by each university. Very few high school students transfer to the technical and vocational area because staying in higher education, passing the joint entrance examination, entering a university and getting a bachelor's degree is expected by most parents and by society. Thus, students who have finished five years of professional school studies can take transferring school tests held by each university or by joint universities and continue their study as third-year university students.

With the development of society and the economy, the expectations and requirements for professional skills and knowledge have become higher and higher. In response to these developments, the Ministry of Education started to upgrade the professional schools. The first 'technological institute' (college level) was founded in Taipei in 1974 and was upgraded to the first 'university of science and technology' (university level) in 1997. More and more professional schools successfully upgraded to institutes of technology after this and then upgraded again to universities of technology. These upgraded institutes and universities of technology continue to provide technical and vocational education. In the late 1990s, an 'educational revolution' movement was started by parent groups and human rights groups and then continued by the Ministry of Education and researchers. This movement aimed to release the pressure on students, free them from examinations and offer more opportunities to study in universities. Thus, the Ministry of Education founded new national universities and national universities of technology, and largely allowed the establishment of new private colleges, universities and institutes. Also, in recent years, the Ministry of Education changed its policy, allowing high school graduates to enter the universities/institutes of technology and also allowing vocational high school graduates to enter universities. The boundary between higher education and technical and vocational education has thus become much weaker. However, most students in institutes of technologies (including the one in which I collected data) still come from vocational high schools.

During the three and a half years when I worked in the college where I collected the data for the study, the college and department offered courses for several educational systems (represented in grey zones in the figure below), including

1. five years of professional school for 15- to 20-year-old students after three years of compulsory junior high school (the ages of students mentioned here and below are for most students in normal situations, but may vary for some reasons, such as illness, accidents or returning to campus after a few years of work);
2. two years of professional school for 18- to 20-year-old students after three years of industrial or commercial vocational high schools (vocational high schools were not offered by the institute of my data collection)

and started new courses in the past decade:

3. following the first or second educational system introduced above, the continuing two years in a technical institute for 20- to 22-year-old students who have finished five years or two years of professional school;
4. four years in a technical institute for 18- to 22-year-old students after three years of vocational high school;
5. two-year master’s degree program for experienced students in electrical and mechanical areas.

The educational systems in Taiwan in the past two decades is described in Figure 3.1.

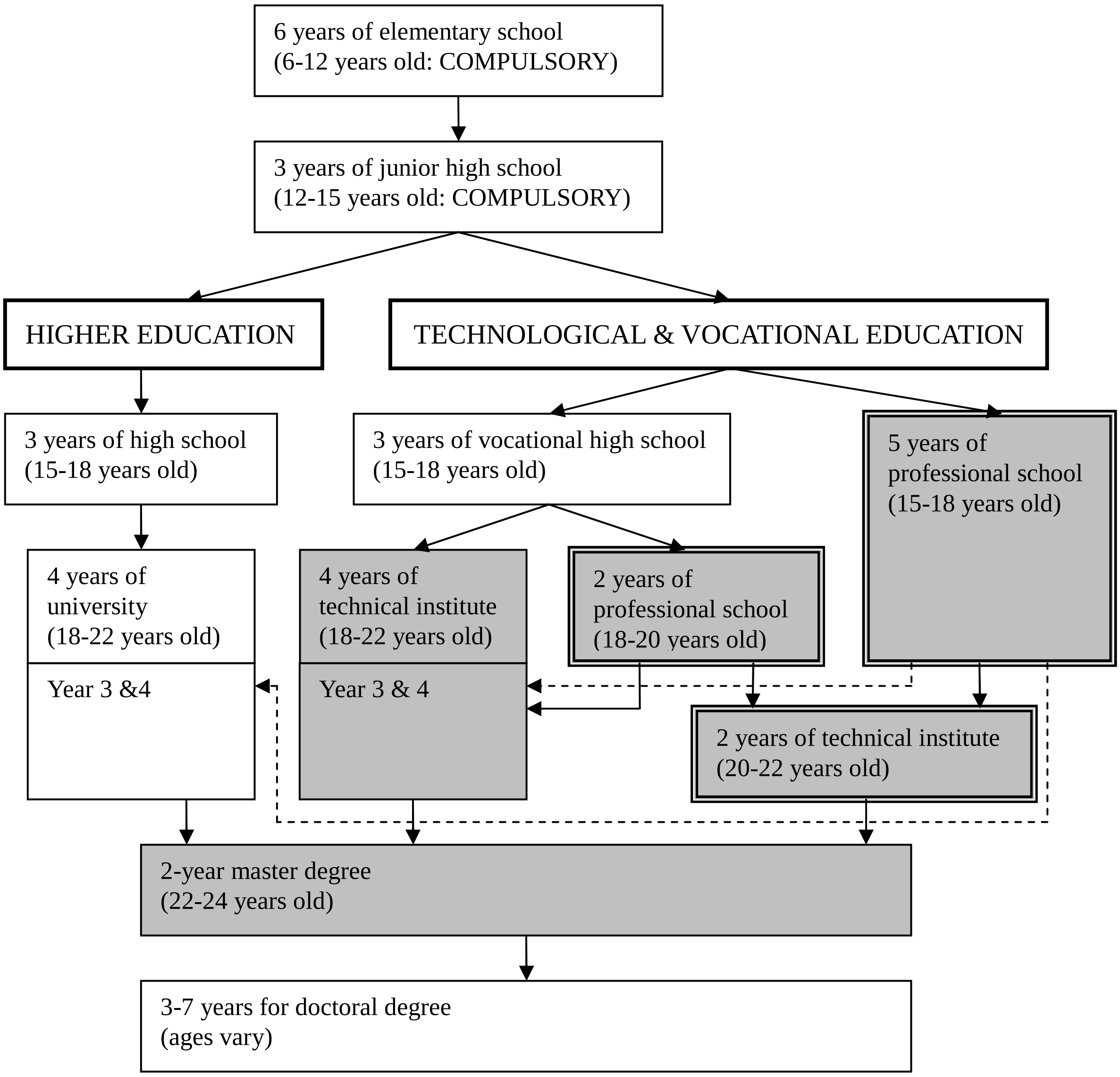


Figure 3.1 Educational systems in Taiwan in the recent two decades

* *The grey zones represent educational systems offered by the Institute of Technology in which I collected data; the educational systems with double-line borders are rarely offered now.*

** *'Technical institute' used in this figure equals the status of 'technical university'. The term "technical university" is not used because this Institute of Technology did not upgrade yet.*

*** *The two choices for five-year professional school graduates are represented in dotted lines in order to avoid confusion with other lines.*

With the new choice of studying in a technical institute for four years, fewer and fewer vocational high school graduates have selected two years of professional school and two years at the technical institute, since that would require one more entry test. Nowadays, junior high school students who choose to study in technical and vocational areas prefer three years of vocational school to five years of professional schools since they can continue studying in the four-year technical institute easily. For the two-year or five-year professional school graduates, entering two years of professional school or transferring into the third year of four years at a technical institute are both available options. However, with the declining numbers of students going to professional schools, the two-year technical institutes are losing their market, and as a result their courses are rarely offered now.

When I was still working for this college, it was transformed from a professional school to an Institute of Technology, and the courses offered for professional schools were gradually replaced by courses for technical institutes. Statistics was taught in the third year of the five-year professional school, the second year of the two-year professional school and the second year of the four-year technical institute, but it was not taught in the two years at the institute since students should have learned this subject in professional school. During my time as a lecturer in the college, I taught statistics to several groups of the two-year professional school and the four-year technical institute in both day school and evening school and to two groups of the five-year professional school in day school.

Before I returned to the Institute of Technology to collect data in 2009, I realised that the department offered courses for the four years at the technical institute in the day school, and courses for the four-year technical institute, two-year professional school and two-year technical institute in evening school, but student numbers for these courses are declining too. Moreover, most of the students in evening school have a job, and their time on campus (five hours per night) is always occupied by classes, so they might not have time or energy to be interviewed. Therefore, I decided to collect data in the combined groups in day school.

This chapter has introduced the educational systems in Taiwan and the background of the Institute of Technology and the participants. In the next chapter, I will talk about the methods and methodology of this study.

CHAPTER 4

METHODS AND METHODOLOGY

4.1 Site and participants of the study

The data for this study were collected in a countryside private Institute of Technology (college level) in a southern county in Taiwan. I was employed as a lecturer in the Department of Business Administration in this college for three and a half years (totalling seven semesters) before I started my Ph.D. course. During these three and a half years, I taught management, mathematics for business, statistics and statistical software (SPSS). I was assigned to teach statistics to single group or multiple groups in each semester. The statistics courses were assigned to several other lecturers after I left. I was allowed to be on unpaid leave and began my Ph.D. study in 2007. I am still an employee of the department now.

When I was in the planning stage of my research and decided to collect data in this college in the second semester in spring 2009, I contacted the department office by telephone to ask for permission to collect data from classrooms and students in the department. I explained the plan, focus and goal of my research and the benefit which might be brought by my research results to the statistics courses and students in the department to the Head of the Department, who allowed me to carry out data collection. I was also informed about some changes at the college and department, the situation, the name of the teacher of the statistics course in the current semester and the number of students (forty-nine year-two students and two year-four students) in the course.

As mentioned in Section 3.1, I decided to collect data in the day school. For the students doing four years at the Institute of Technology in the Department of Business Administration, statistics is taught in both semesters in the second academic year. There were two groups of second-year students, and they were combined in several subjects, including statistics. Normally an annual course was assigned to one teacher, but the teacher of the statistics course in this academic year changed at the beginning of the second-half semester because the previous teacher had been transferred to another department. The teacher in charge of the second semester was in the Department of General Education, and luckily we were familiar with each other when I was working for the institute. This reduced the difficulty for me to get permission for data collection in his class, but in accordance with UEA's requirements of field work for research projects involving human participants (and the BERA guidelines), a consent form (see Appendix 4.1.) was given to the teacher.

I contacted the teacher and introduced my study by telephone, and we met before I started my data collection. We had a short discussion when I returned to Taiwan in the seventh week of the second semester. The plan, focus and goal of my study were explained to the teacher, and he agreed to be observed in class and interviewed after class and signed the consent form. After viewing the students' tight schedules, I decided to include ten participants because I believed that this would be a large enough sample given the cohort size, but small enough to allow for enough time to prepare and organise clinical interviews, analyse their responses and adjust the design of the next clinical interviews each time. Therefore, I told the teacher that I would like to include ten students in the study for clinical interviews, and the teacher promised to help me contact them. The statistics course was held from 11:10 am to 12:00 pm and from 1:00 pm to 2:50 pm with ten minutes break at 1:50 pm on every Tuesday, and my classroom observation was made on every Tuesday between 14th April (week 8) and 16th June (week 17), except for 21st April due to the mid-term.

The teacher was in the role of “gate-keeper” (Miller & Bell, 2002, p.53) of the data collection in two ways. Besides allowing me to observe in the classroom, he also introduced me to the students of the statistics course in my first classroom observation. He told the students that the focus of my classroom observation was on his teaching as well as their learning and asked them to ignore my presence in the class. During the break of the first classroom observation (12:00 pm 14th April), the teacher asked me to wait outside the classroom, and he then introduced me to ten students (five students from each group). The teacher was familiar with the students and knew their mathematical abilities and learning attitudes. He told me, in private, that the cohort was a mix cohort in terms of achievement and that some students did very well while others did not achieve as well. He also believed that the ten students he suggested might be representative of the spread of abilities in the class.

The ten students and I went to the coffee table and several chairs just outside the classroom, and I introduced myself, my background, my experience, and my interest in the study: how the students solved questions in statistical inference and the role of visualisation. I also gave them the consent forms (see Appendix 4.2.). I did not ask them to sign the consent forms immediately but let them take the consent forms home and consider carefully before joining my research. They all agreed to participate and handed in the signed consent forms to the teacher in the mid-term week.

4.2 Ethical issues

Before beginning with my data collection, I considered ethical issues and included them in the consent form for the teacher and students. Foster (1996) refers to guidelines in

BERA (1992) in suggesting researchers build an honest and open relationship with participants. He explains the idea of informed consent:

> …all potential participants in a research project should be able to agree or refuse to participate on the basis of full information about the research (1996, p.105).

And the BERA (2004) guidelines state:

> Researchers must take the steps necessary to ensure that all participants in the research understand the process in which they are to be engaged, including why their participation is necessary, how it will be used and how and to whom it will be reported (BERA, 2004, p.6).

Clough and Nutbrown (2007) also suggest that research should meet ethical standards that

> provide the best possible protection for researchers and their participants;
> ensure that data are collected within informed consent of participants; and
> protect participants' personal details, identities and well-being (p.173)

Thus, my ethical considerations of doing this research were as follows:

Anonymity: The most important thing is to protect the privacy of all participants and to avoid their exposure to any others, even their classmates. Mertens (1998) explains anonymity as "no uniquely identifying information is attached to the data" and "no one can trace back to the individual providing it" (p.279). Thus, the real name, similar name or nickname of each participant who takes part in this research will not appear in my data, my thesis or any other published/unpublished paper, in order to prevent giving any clues that can be traced back to the participants. Alternatively, I will use codes S1, S2,… and S10 to represent each participant. The name of the institute will not be mentioned, either.

Confidentiality: As part of the consenting procedure, participants in this study were provided with a written commitment that all data would be kept confidential and that the only purpose of the data collected from them would be for academic research and my doctoral thesis. The information sheet also explained that no one but my supervisors and I would see the original data. The identity of the participants and all the data collected in this study, such as video/audio recordings in class, group discussion and interviews, students' homework and answer sheets in examinations and class materials would be kept totally confidential. Finally, each participant's data would only be viewed and confirmed by him/herself and me and would not be released to any other person.

Gate-keeper: For the students, the teacher played the role of the gate-keeper who is "in a position to 'permit' access to others for the purpose of interviewing" (Miller & Bell, 2002, p.55). I explained to the teacher my interest in the study and my role in/after class as a researcher. I would not judge his teaching or any other aspect in the statistics courses. I needed the teacher's permission to collect data in his class. I also needed him to introduce me and my research to the students at the beginning of class observation, and to recommend students to participate.

Power issues: Before I started collecting data in the first classroom observation, the teacher introduced me and my task to all students in the statistics class. Although I had gone to UEA for doctoral study two years earlier, I was a lecturer on unpaid leave and still a member in the department. Some students might find out who I was and feel that I came to judge them, which raised some power issues. For example, they might be afraid of the consequence of joining my research, such as a grade being decreased by me or a negative evaluation provided to the statistics teacher, their mentor or even the department. On the other hand, they might consider the consequence of not participating: would I mark them or influence their score? They might also worry about what I would do if they performed badly in the clinical interviews or withdrew from the study. I therefore clarified my role as just a researcher to the participants.

Role of researcher: I explained to the students that their participation in this research was only for academic purposes. To avoid students' over-expectation of being a participant, I made clear that I would not be able to change their grades or have any influence on their marks, except for what they learned in the study. However, I offered to give a financial reward to the participants at the end of data collection to thank them for their contribution and time spent on the study.

Freedom to participate and withdraw: The willing students and teacher in the statistics courses in the department could participate in this research without any requirement or condition. Also, they could leave the research at any time without fear of any consequence if they did not want to continue. The BERA (2004) guidelines remind the researchers to notice, avoid and reduce participants' possible distress and discomfort in the research process. I gave the students my email address and mobile phone number so they could contact me when they had any thoughts or ideas about my research. With the teacher's permission, I also gave the teacher's e-mail address and mobile phone number to the students, so they could contact the teacher if they had any problem or felt uncomfortable with my research.

Consent of teacher and students: When the teacher consented to participate, he was guaranteed that his privacy would be protected as much as possible. I also promised his safety of joining/not joining/leaving this research and maintained his right to check the accuracy of collected data, delete unwanted data and view the research results. Furthermore, I gave consent to the students to explain what I would collect and observe in the class, what I would collect after class (such as their homework and answer sheets from examinations), what area of questions I would address and what equipment (e.g. video or audio recorder) I would use in the interviews and for my role as researcher in the study. There were more than fifty students in the statistics course, and I wanted to find ten participants for clinical interviews. Most interviewees formed focus groups because they preferred to discuss matters with classmates. Throughout the research process, consent was "ongoing and renegotiated" (Miller and Bell, 2002, p.67).

In consideration of the ethical issues, I submitted an ethics form (Appendix 4.3.) to the Research Ethics Committee at the School of Education and Lifelong Learning. The study was approved by the Chair of the Committee via e-mail (see Appendix 4.4.) when I was away from the UK.

4.3 Methods for data collection

The main data for this study were collected by classroom observation, teacher interviews and student interviews. The preparatory paper and the teacher's solutions, along with the questions in the mid-term and final examinations were also collected (see Appendix 4.5. - Appendix 4.10.). Presmeg (2006) suggests that classroom observation and clinical interviews can be used to go deep in visualisation research in mathematics education. These approaches are explained as follows.

4.3.1 Observation

As a popular research method in educational research, observation was given a working definition by Clough and Nutbrown (2007), as

> …simply ‘looking’ - looking critically, looking openly, looking sometimes knowing what we are looking for, looking for evidence, looking to be persuaded, looking for information (p.50).

My data collection began with classroom observation. The initial focus was the teacher’s teaching, question-solving and use of visualisation. Whenever unexpected phenomena were found useful and relevant to my study, they would also be included in the field notes. The classroom observations were the basis from which to design the questions for the teacher and student interviews.

Foster (1996) suggests several advantages of observation, such as providing detailed information about school life which cannot be acquired through other approaches. Cohen *et al.* (2000) argue that observations enable researchers to “collect live data from live situations” (p.305). In addition, the observation method prevents researchers from relying solely on what is said in interviews or what is written in documents. In other words, through observation researchers can record first-hand information by seeing it when it happens, and it therefore seems to be more accurate than relying on what is said by others at a later time. Furthermore, observers may see what participants cannot see in a familiar school environment. For those who agreed to be interviewed, observation can complement information missed or unanswered in the interviews (Foster, 1996).

I attended the statistics class from the eighth week to the seventeenth week, except for the ninth week due to the mid-term in the second semester. My focus was the teacher’s teaching, question-solving and use of visualisation. With his permission, two audio-recorders were used throughout all class hours. The batteries were charged beforehand to ensure that they functioned well in class. I watched and listened to the teacher, transcribed whatever he wrote on the blackboard and made field-notes. The teacher did not wish to be recorded on video at all times, but he understood that visualisation was my main concern, so he allowed me to use a camera or video-camera when there were incidents involving visualisation (graph, table or other forms of visual representation). Therefore, there is a somewhat limited amount of pictures and video-recordings of the class. I also collected the teacher's materials and textbook. He did not give any handouts during the eleven weeks of data collection, except for two practice question-sheets before the mid-tem and final exams. He taught the course according to the textbook (Kuo and Shi, 2006) by introducing theories and solving example questions in a traditional way. Students sat and listened, and they rarely asked questions. The data observed in class were analysed and then combined with the findings from teacher and student interviews and the analysis of the teacher's materials and students' products to design questions for the teacher interviews and student interviews (for the details, see Chapter 6).

4.3.2 Interviewing

Bryman and Burgess (1999) suggest that interviews are one of the most important methods for data collection. According to his experience that children's unexpected responses (such as wrong solutions) are often interesting and valuable but cannot be measured by traditional standard tests, Piaget created the clinical interview as

> a flexible method of questioning intended to explore the richness of children's thought, to capture its fundamental activities, and to establish the child's cognitive competence (Ginsburg, 1981, p.4).

Ginsburg (1997) explains that a clinical interview is not simply an ordinary conversation but rather a conversation with a purpose: to enter the mind. Berg (2000) distinguished three major categories of qualitative interview: structured formal standardised interview, non-directed informal interview, and semi-structured (semi-standardised) interview.

Lindberg and Rosenqvist (2003) suggest that the face-to-face and semi-structured interviews are among the most common approaches for interviewing in qualitative research. This kind of interview involves the implementation of a number of predetermined questions and/or special topics. These questions are typically asked to the interviewee in a systematic and consistent order, but the interviewers can ask additional questions according to the interviewee's responses. By using open-ended questions and both planned and unplanned prompts, the interviewers are permitted to explore far beyond the answers to their prepared and standardized questions (Berg, 2000). Khera *et al*. (2001) indicated that semi-structured interviews allow the respondents to determine the direction and content of the interview within a broader framework provided by the interviewer (Chang, 2004).

Therefore, in this study, I decided to carry out semi-structured clinical interviews with the teacher and the students, because this method enables me to not only obtain information of my interest in the designed questions but also collect and respond to the unexpected answers from the interviewees, which may be valuable for my study. For convenience, I will simply call the interviews undertaken in the study clinical interviews since the main purpose is to investigate the participants' thought processes when solving statistics questions.

4.3.2.1 Interviewing the teacher

The interviews with the teacher were initially arranged in the ninth, twelfth, fifteenth, and eighteenth weeks of the second semester, but eventually the teacher was only interviewed in the ninth and sixteenth weeks because he was extremely busy on his Ph.D. course.

In the first interview, the teacher explained that he tested the same questions in the mid-term because most of the students were not good at studying and he did not want to see them fail because of their low scores. He also explained the advantages and his interpretation of visualisation, the reasons for choosing a particular table of normal distribution and for asking students to memorise several z scores which could be found in the table, and his experience of teaching standardisation. His responses were referred to when designing questions in the first student interviews.

In the second interview, the teacher was exhausted because of his preparation for the examinations in his Ph.D. course, and he was not in a position to answer the questions. Unfortunately the data was not substantive enough, and he could not make the third interview because of the tight schedule of his Ph.D. course and examinations.

4.3.2.2 Interviewing the students

The statistics course was taught to two combined second-year groups, and there were 49 students who were 19 to 20 years old, with two students who were 22 years old. I did not know any students in that class, so I appreciated the teacher introducing me to them and recommending ten volunteer students to participate. The ten participating students came from two groups. In order to protect their identities, their names will not be mentioned in this study but will be replaced by codes. S1, S2…and S5 were from one group, and S6, S7… and S10 were from the other group. Most participants felt nervous and preferred to be interviewed with others, so I suggested to them that they form focus groups in order to help ease their tension and allow them to share their ideas and thoughts (Madriz, 2003). Madriz (1998) argues that focus groups enable the group members to exchange, verify and confirm their experience with others. Brown and Liebling (2005) add that sharing ideas is a good method for exploring misconceptions, and that it is better to let the person who knows the most to say the least and let the person who has difficulty with the exercise questions to talk about their thinking. Therefore, nine of the participating students formed three focus groups of S1 and S2, S3 and S4, and S6, S7…and S10. These students all requested to be interviewed with familiar classmate(s) when they initially decided to participate in this study. Only S5 agreed to be interviewed individually.

Ginsburg (1981) explains that the clinical method enables the interviewers to treat each interviewee flexibly and differently for three goals:

> the discovery of cognitive activities (structures, processes, thought patterns, etc), the identification of cognitive activities, and the evaluation of levels of competence (Ginsburg, 1981, p.5).

For the goal of discovery, he argues that it does not make sense to define mathematical thinking in the early stage of studies, and researchers should first carefully observe, explore and try to discover. Thus, how participants think of using the tables of distribution is formed and refined after clinical examination. Piaget (1929) explains that the clinical examination depends on direct observation. Ginsburg suggests that the clinical interview should be an “open ended” (Ginsburg, 1981, p.6) task and that the interviewer should possess a contingent attitude to obtain as many reflections as possible, while simultaneously working to avoid biasing the interviewees’ responses.

For the goal of identification, I believe that clinical interviews enable me to explore students' individual steps and cognitive processes of how they use the table of distribution with/without considering the graph. There are three sub-goals necessary to achieve this. First, rich verbalisation is required because "[t]he complexity of knowledge may not be revealed by simple response" (Ginsburg, 1981, p.7). Second, misinterpretation should be avoided carefully by checking the reports and clarifying ambiguous statements. Third, alternative hypotheses concerning the underlying processes should be tested.

For the investigation of levels of competence, the interviewer needs to evaluate whether the interviewee is motivated to do the task, whether he/she realises what the question really is and whether he/she is confident about his/her answer.

Before using the clinical method, the purpose for research must be clear because "[t]here is, in effect, no one clinical method; there are three, each designed for a different research purpose" (Ginsburg, 1981, p.10). In the interviews, I let the students do some exercises, and I observed how they developed and took advantage of visualisation in thinking processes and problem-solving rather than only looking at the results, since process-tracing methods provide more information than input-output methods (Pugalee, 2004).

I also discussed with the participants their products in the interviews. Two audio-recorders were used through all the interviewing processes, and video-cameras were used when necessary. The audio and video recordings were transcribed and analysed as soon as possible after the interviews. They were first interviewed in the tenth and eleventh weeks, and they were interviewed again in the fourteenth and fifteenth weeks. Each interview took seventy to ninety minutes. S1, S2… to S5 did not prepare well for the second interview, so they did not solve many questions, especially the hypothesis testing question. Thus, they were asked and agreed to have a third interview together in the eighteenth week after the final examination. S6, S7… to S10 did not attend the third interview because they did not have enough time after the final examination for personal reasons.

4.3.3 Examining students' products

A time-table of data collection is displayed in Table 4.1.

Event \ Week	8	9	10	11	12	13	14	15	16	17	18
Content	Review		Estimation		Hypothesis testing					Review	
Examination		Mid-term									Final
Answer sheet			V								V
Classroom observation, notes & audio/video record	V		V	V	V	V	V	V	V	V	
Teacher interview		V							V		
Student interview I			S1 -S5	S6 -S10							
Student interview II							S1 -S5	S6 -S10			
Student interview III											S1 -S5

Table 4.1. Time-table of data collection.

Students' products are a small part of the data since the teacher did not assign homework during the eleven weeks of data collection, as he had found that very few of them did homework on their own, and the participating students' notes were the same as what the teacher had written on the blackboard. Therefore, the only products of the students collected were their answer sheets to the mid-term and final examinations. The ten participants' scores in the mid-term and final exams are listed in Table 4.2.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Mid-term	88	69	100	96	84	100	98	98	93	100
Final	78	79	95	85	87	100	95	94	90	97

Table 4.2. The ten participants' scores in the mid-term and final examinations.

The participants' answer sheets of the mid-term and final examinations were photocopied with their permission, but the answer sheets might not represent how the participating students understood the questions and how they attempted to solve the questions, since the solutions of most questions had been revealed by the teacher. This may explain why some participants' scores did not match their performances in the exercise questions (which will be discussed in Chapters 6, 7, 8, 9 and 10).

I decided not to take into account the students' responses to the mid-term and final exams in the study, but I will analyse the questions in the mid-term and final exams to find out the question types the teacher set and, according to the analysis, design exercise questions to reexamine the students' learning results and to explore the students' learning strategies and difficulties.

4.4 Analytical Tool

According to Charmaz (2005), grounded theory methods provide useful tools for analysing process beyond description and enable researchers to stay close to the interested area. Also, these methods help the researcher to

> develop an integrated set of theoretical concepts from their empirical materials that not only synthesize and interpret, but also show the processual relationships (Charmaz, 2005, p.508).

Grounded theory methodology is a process of “constant comparison’’ (Glaser and Strauss, 1967, cited in Harry *et al.*, 2004, p.5) in which researchers “[move] back and forth’’ (Harry *et al.*, 2004, p.5) among the data and resulting theory. Charmaz (2005) explains that grounded theory methods consist of simultaneous data collection and analysis. She argues that these methods encourage researchers to begin analysis early and help focus further data collection; the focused data can then be used to refine the emerging analyses. This follows Glaser’s (2002) argument that “conceptual categories arise through our interpretations *of* data, rather than emanating *from* them or from our methodological practices” (Glaser, 2002, cited in Charmaz, 2005, p.509), adding that the theoretical analyses are “interpretive renderings of a reality, not objective reportings of it” (Charmaz, 2005, p.510). However, Deutscher *et al.* (1993) remind interviewers that what people say may not be what they do, and they are “not always in concordance” (1993, p.17) since interviewers’ interpretations may be influenced by the current situation and their theoretical background. Grounded theory enables researchers to “compare the data with data, data with categories, and category with category” (Charmaz, 2005, p.517).

During and after the eleven weeks of data collection, classroom field notes, audio and video recordings made in the classes and teacher and student interviews, students’ products, including mid-term and final examinations and exercises in the interviews, were collected, transcribed and then analysed. The data collected in the interviews were analysed according to the principles of open coding (Corbin and Strauss, 2008), revised and regrouped. Corbin and Strauss (2008) state that

> …analysis involves examining a substance and its components in order to determine their properties and functions, then using the acquired knowledge to make inferences about the whole… Analysis is a process of examining something in order to find out what it is and how it works (2008, pp. 45-46).

Some students did not say much, and many of their answers were short and simple. However, Corbin and Strauss (2008) remind us that "there are no simple explanations for things" (2008, p.8) and that we should explore students' thoughts through their answers and responses. The way we give the collected data significant meaning is then called analysis. They suggest that microanalysis, which is like using a microscope to examine each piece of data, can be used in the beginning stage to break into and make sense of the data. When doing this, researchers should observe carefully, pay attention to details and keep an open mind. They explain that microanalysis has four components:

1. generating probabilities;
2. checking out the probabilities against data when doing 1;
3. discarding the irrelevant; and
4. revising interpretation when needed (Corbin and Strauss, 2008, p.60).

Corbin and Strauss (2008) distinguish microanalysis and general analysis: "microanalysis looks at detail, general analysis steps back and looks at data from a broader perspective" (2008, p.6) of what all the data tells us. Microanalysis and general analysis complement and supplement each other. General analysis is easier to do, but microanalysis prevents researchers from skipping and missing any possible interpretation.

Corbin and Strauss (2008) suggest several types of analytic tools to interact with and make sense out of data. Analytic tools can prevent researchers from being constrained by standard thinking approaches about phenomena and keep them away from the technical literature and personal experience which may prevent them from seeing new possibilities in data. The most basic and useful analytic tools they recommend are questioning and making comparisons. They explain that questioning (asking questions) is “fundamental to analysis” (2008, p.69) and that good questions enable “discovery of new knowledge” (2008, p.69). Asking questions helps researchers go through the initial block of not knowing where to start. Also, when asking questions about data, more questions will be generated, enabling the analyst to “probe” (2008, p.69) deeper into the data and avoid shallow and uninteresting findings.

Generally speaking, there are two types of comparison: “constant comparisons” and “theoretical comparisons” (Corbin and Strauss, 2008, p.73). Corbin and Strauss (2008) explain that constant comparisons are used to compare each incident for similarities and differences, after which similar incidents can be conceptually grouped together. By doing this, researchers will be able to identify properties and dimensions of one category/theme and distinguish it from others. Moreover, comparison within one category/theme helps in exploring more of its different properties and dimensions.

When researchers are unable to define an incident according to its properties and dimensions, or are unsure how to classify it, they should turn to theoretical comparisons which are designed to help researchers obtain a definition or understanding of phenomena by looking at the property and dimension levels. When doing this, the researchers can draw upon what they know to understand what they do not know, discover the similarities and differences between each object, define them and come to know something through their properties and dimensions.

Making comparisons requires properties which are then used to examine the incidents of data, and specific incidents used in theoretical comparisons can be derived from the literature and experience. However, theoretical comparisons would be unnecessary if properties were evident in the data, as Corbin and Strauss (2008) make clear:

> …it is not that we use experience or literature as data, but that we use the properties and dimensions derived from the comparative incident to examine the data in front of us… it is not the specifics of an experience that are relevant but the concepts and understanding that we derive from them (2008, pp.75-76).

Using comparison as an analytic tool forces researchers to think about the property/dimension level rather than the specifics or raw data level, which moves them from a descriptive level to a more abstract level.

4.5 Coding

Charmaz (2005) suggests that coding is the first step in the data analysis process, as it offers researchers “analytic scaffoldings on which to build” (p.517). Because researchers stay close to their data, they can define “new leads from them and gaps in them” (p.517). Coding also helps researchers see their assumptions and reconsider why and how they develop such codes.

Coding, which means to name codes, is a necessary step in analysis, and it takes and raises raw data to a conceptual level. Corbin and Strauss (2008) suggest that coding is like “mining” (p.66) the data: digging through the surface to discover the hidden treasures. They stress that coding is not merely paragraphing, noting concepts in the margin of the field notes or making a list of codes as in a computer program, but rather interacting with data, deriving concepts to stand for data and developing concepts "in terms of their properties and dimension" (p.66).

Here I explain how I summarised and distinguished the participants' difficulties during the eleven weeks of data collection.

Strauss and Corbin (1998) suggest open coding, which means to name events and actions in the data as the first step of the development of such a method. In their study, Harry *et al.* (2004) began with open coding of collected data, then moved through categories and themes to theory development, and finally drew a data analysis map to express the contents, levels and relationships. They also pointed out their view that the limitation of their code-mapping model was its impression of linearity.

The participants' solving steps in all exercise questions in my study are also represented in a linear model, but it is difficult to represent all events in the order of time. For example, some processes were revisited by individuals, while some of the necessary processes might be omitted by individuals. The repeated steps were linked by two separate opposite directional arrows rather than a two-headed arrow in order to indicate that one step might influence the other and that the arrow on the left-hand side occurred earlier than that on the right-hand side. The revisiting steps would not stop until the transferred equation was regarded by the individual as suitable for his/her selected table of normal distribution. For instance, the interpreted z score and the relation sign would be compared to the format of the chosen table of normal distribution to determine whether a further transferring was needed or not. If transferring was not needed (which means the relation signs were the same and the z score was included in the table), the given z score would be the z score for searching; if transferring was needed and indeed applied, the new relation sign and the generated z score in the transferred equation had to be compared to the format of the table again. In fact, some participants did follow such a route, and how they solved the questions will be compared and analysed; some participants did not transfer the equation or did not get the correct transferred equation, and their problem and omission were also pointed out and analysed. By doing this, the revised solving steps are grounded in their solutions.

In this chapter, I introduced the site of data collection, the participants and the ethical consideration of this study, and I also explained my methods of collecting data. Then, I explained how I analysed and categorised the data. In the next chapter, I will discuss some issues observed in class and data which relate to my study.

CHAPTER 5

GENERAL OBSERVATION

5.1 Content of the course

The data were collected between the eighth week and the eighteenth week of the second semester in 2009, and my first classroom observation was made on the eighth week, when the teacher gave practice questions in preparation for the mid-term test in the ninth week. The questions involved content taught in the previous seven weeks as well as some relevant basic ideas taught in the first semester. These questions included two main parts. The first part asked students to 'fill in' the formula according to the question, such as

1.1 formula of standardisation;

1.2 mean and standard deviation of sample proportion; and

1.3 sample mean and sample standard deviation of normally distributed variable.

The other part tested students' ability of application about

2.1 finding the probability of normally distributed variable after standardisation;

2.2 formula of standardisation, and finding the probability of normally distributed variable after standardisation;

2.3 formula of sample proportion, and finding the probability of sample proportion;

2.4 finding the 95% confidence interval of sample proportion; and

2.5 finding the 90% confidence interval of population mean.

The teacher used the textbook *Statistics* written in Chinese by Kuo and Shi (2006). The content taught during the first seven weeks in the second semester were

Chapter 8: Estimation:

Point estimation and interval estimation of population mean (μ);

Point estimation and interval estimation of population proportion (p); and

Point estimation and interval estimation of population variance (σ^2).

Some knowledge relevant to Chapter 8 was included in the previous two chapters and taught by another teacher in the first semester. This content was also tested in the mid-term, such as

Chapter 6: Continuous probability distribution:

Normal distribution (N);

Standardisation; and

Standard normal distribution (Z).

Chapter 7: Sample distribution:

Mean and standard deviation of sample mean ($E(\overline{X}), \sigma_{\overline{X}}$);

Mean and standard deviation of sample proportion ($E(\hat{p}), \sigma_{\hat{p}}$); and

Sample distributions of small sample: χ^2 distribution, t distribution and F distribution.

However, some ideas such as F distribution were skipped in the first semester, and the teacher in the second semester did not introduce the concept, instead skipping some content (such as testing variance) relevant to F distribution.

In the beginning of the tenth week of this statistics course, the teacher gave the marked answer sheet to the students, let them check the correctness of their scores and gathered all answer sheets back. He did not explain the questions again since most questions were the same as the practice questions. He started teaching new content until the sixteenth week. Again, he gave practice questions in the seventeenth week. The content taught between the tenth week and sixteenth week were:

Chapter 8: Estimation:

- Point estimation and interval estimation of difference between two population means ($\mu_1 - \mu_2$);
- Point estimation and interval estimation of difference between two population proportions ($p_1 - p_2$); and
- Point estimation and interval estimation of ratio of two population variances (σ_1^2 / σ_2^2). The teacher simply explained the ideas and skipped much of the content in this section when students told him that F distribution had not been taught previously.

Chapter 9: Hypothesis testing:

- Ideas and types of hypothesis: null hypothesis (H_0) & alternative hypothesis (H_1), Critical region (CR) & accept region (AR);
- Ideas and types of test: two-tailed test, right-tailed test & left-tailed test;
- Type I error (α) & type II error (β);
- Process & methods of hypothesis testing;

Testing the population mean (μ);

Testing the population proportion (p);

Testing the population variance (σ^2): the teacher skipped this section because of its difficulty;

Testing the difference between two means ($\mu_1 - \mu_2$);

Testing the difference between two proportions ($p_1 - p_2$);

Testing the difference between two variances (σ_1^2 / σ_2^2): this section was also skipped.

The teacher gave practice questions in the seventeenth week as preparation for the final examination. At the front of the questions, some z scores, t scores and a complicated formula of pooled sample variance S_p^2 which might be used in the solving processes were attached as follows:

$$Z_{0.975} = 1.96, Z_{0.99} = 2.325, Z_{0.995} = 2.575,$$
$$t_{0.01}(10) = 2.764, t_{0.01}(11) = 2.718, t_{0.1}(11) = 1.363$$
$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Therefore, students did not refer to the tables of normal distribution and the table of t distribution to find the needed z scores and t scores by themselves. The questions tested in the final examination also included two parts. In the first part, students were asked to fill in the formula of

1.1 test statistic of mean of normal population with a known population standard deviation;

1.2 test statistic of mean of normal population with unknown population standard deviation, small samples;

1.3 test statistic of difference between two means of independent normal population with known population standard deviation;

1.4 test statistic of difference between two means of independent normal population with unknown and equal population standard deviation, small sample;

1.5 test statistic of difference between two population proportions.

The second part included six application questions. Students should not only find the correct formula, but also enter the corresponding numbers and calculate to get the results. The six questions included:

2.1 to find the 99% confidence interval of difference between two means of independent normal populations with unknown and equal population standard deviation;

2.2 to find the 95% confidence interval of difference between two means of independent normal populations with unknown and equal population standard deviation, small sample;

2.3 to find the 99% confidence interval of difference between two proportions;

The last three questions were placed on the second page of the question sheet, and the information attached at the beginning was attached at the end of question sheet again. The three questions were:

2.4 two-tailed Z test for a mean & left-tailed Z test for a mean;

2.5 two-tailed Z test for difference between two means; and

2.6 two-tailed t test for difference between two means of independent normal population with unknown and equal population variance, small sample.

During the eleven weeks of data collection, the taught content consisted of estimation (before the mid-term) and hypothesis testing (after the mid-term). The mid-term test was taken in the ninth week, while the final examination was in the eighteenth week. The teacher did not introduce new content in the eighth and seventeenth weeks, but rather gave questions to the students for practice. Most (nearly 90%) of the questions in the practices were the same as the questions in the mid-term and final examinations because the teacher was worried about the students' performance and did not want them to fail. Thus, the teacher let the students practise to get familiar with the questions.

5.2 Students' application of the table(s) of normal distribution

Up to the first interview, participating students' use of tables of normal distribution in class was simple: to find the given z score outside the tables of normal distribution and its corresponding probability inside those tables. Using the tables of normal distribution in this way will still be mentioned in later contents, and here I name this 'outside-in'. Questions in this format were tested in Q1.1-Q1.6 in the first interview. Using the tables of normal distribution in such a manner can be expressed by Figure 5.1.

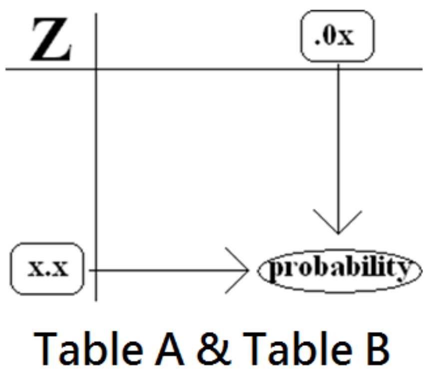


Figure 5.1. Graphical representation of using the tables of normal distribution in the direction of 'outside-in'.

Although there has been much research concerning different topics in statistics education, the real situation of how students use the tables of distribution, what difficulties they encounter and why they experience problems seems not to have been addressed in previous research. Therefore, I feel it necessary to explore such issues. Dreyfus (2006) refers to a suggestion made at PME 28 that

> mathematics education at present needs models to fit specific research questions rather than big, unified theories (2006, p.80).

Using the tables of normal distribution to solve Q1.1-Q1.6 involves many items, such as the meaning of the questions, the use of the tables, the meaning of elements in the tables, the relations between the elements in each table, the difference between Table A and Table B, the format of Table A and of Table B, graphical representation of the question, etc. Solving a question involves decisions, choices and the chance to select and execute a suitable way (Brown and Liebling, 2005). Bernard and Tall (1997) suggest that it is important to compress information into cognitive units and to connect cognitive units in mathematical thinking. Some students may build and connect more cognitive units than others and solve the question with fewer difficulties. For example, some students in this study drew graphs to help them think of the necessary transferring steps (see Section 5.3) and avoid problems.

Corbin and Strauss (2008) suggest analysing by

> taking data apart, conceptualizing it, and developing those concepts in terms of their properties and dimensions in order to determine what the parts tell us about the whole (2008, p.64).

They add that initial analysis should be more detailed to explore all possibilities for interpretation, and that later analysis should become more general to develop and validate interpretations. Corbin and Strauss (2008) expand Lamott's (1994) idea that asking questions helps deal with the "initial block of not knowing where to start" (2008, p.69). Therefore, when I found that students had different kinds of difficulties in solving the exercise questions in the interviews, I kept asking myself questions, such as how these questions should be solved, how students solved them, what mistakes students made in their solving steps, why they made these mistakes, and how students attempted to overcome difficulties, etc. Legutko (2008) suggests a systematic approach to solve problems: "draw a figure, write formulas, transform them, substitute data, calculate and… get the result" (2008, p.149).

Kirsch *et al.* (1998) developed a framework to evaluate the degree of difficulty of questions by analysing how respondents read the information in the prose, document and quantitative questions. Although the focus of my study was on how the participants solved the questions and in which step(s) the problems occurred, rather than on knowing how difficult/easy the questions were, Kirsch *et al.*'s (1998) process variables provide ideas to deconstruct the exercise questions and some solving steps.

By observation and comparison, I developed a process of solving the questions of using the tables of normal distribution to find a value. The process can be explained as ongoing actions taken in response to problems (Corbin and Strauss, 2008). The solving steps of Q1.1-Q1.4 were first developed. The solving steps of other groups of exercise questions were built on the basis of the solving steps of Q1.1-Q1.4 and revised in terms of the condition of the questions. The initial solving steps of Q1.1-Q1.4 consisted of four stages:

1. **Question**: how the question was represented to question solvers. Unlike Kirsch *et al.*'s (1998) exemplar questions, which require something in sentences whereas given information was stated or implied in other texts, each of the exercise questions was an equation which contains not only the requested information, but also the given information. Kirsch *et al.* (1998) distinguish that

   > *Given information* is information that is known and assumed to be true based on the way a question or directive is stated. *Requested information* in a question or directive is information being sought (1998, p.114).

   Taking Q1.3 '$P(Z > -3.1) = ?$' for example, the question asking the probability of Z being bigger than -3.1, was performed in a symbolic from, and all students were given the same question without further explanation. A symbolic question such as this was represented in an equation with its conditions (i.e. the given information) and the unknown answer (i.e. the required information). The questions were not influenced by the question solvers.

2. **Information**: represents what was given and implied in the question. One question may contain multi-information. In other words, this stage is the decomposition of the question. Taking Q1.3 for instance, the format of $P(Z > -3.1) = ?$ in statistics represents looking for the probability value of Z being bigger than -3.1 in which Z indicates this question involves standard normal distribution, -3.1 is the specific z score, = means that the left-hand side and the right-hand side are the same or equal, and the ? represents the unknown which is to be sought in this question.

   The unknown answers in these questions were all numbers, and the participants were told to find the probabilities. However, the unknowns were not needed when solving these questions and appeared to be invisible before they were obtained. Also, the information given in the question might not be correctly recognised by the students. I observed some of my previous students dealt with such questions on the basis of what they actually interpreted from the question, rather than what they were given by the question in their examination sheets. Thus, I take this issue into account and distinguish what was given and what was interpreted by the question solvers as separate steps.

3. **Interpretation**: represents how the information in the question was actually interpreted by the participants. Kirsch *et al.* (1998) emphasise that the given information and requested information must be first identified. The participants were supposed to make sense of the meanings of what was given and what was required by locating the numbers and signs (i.e. the given information) and the unknown (i.e. the required information). In Q1.a-Q1.4, they were told to find probabilities, but in some other questions it was not explicitly identified what was being sought, and students might thus have some problems in recognising their meanings since the given information and requested information were given in an equation, rather than separate question sentences and information texts in Kirsch *et al.*'s exemplar question. The format of such a question appeared to be a distractor when interpreting the elements in the question. Therefore, this issue must be taken into account.

   Also, what is actually interpreted might differ from what is given, and what is given may not even be noticed. Therefore, students may solve a question with insufficient or incorrect interpretation of the question, such as treating $P(Z > -3.1)$

as P(Z > 3.1) without considering the minus sign, or seeing the question as P(Z < -3.1) without considering the relation sign given in the question. Thus, it is not the given information itself but the interpretation of the information which is taken into account in the individual's solutions. I will discuss the details of how students interpreted the questions in Chapters 6, 7, 8 and 9.

4. **Action**: includes the question solver's work on the interpreted information, such as choice of tables of normal distribution, whether and how the question was transferred to fit the format of the selected table of normal distribution, looking for z score outside the table of normal distribution and its corresponding probability inside the table, and further necessary steps to deal with the found probability if transferring was applied. All the actions to solve the question will be included in this stage, such as choosing the table of distribution, comparing and contrasting the format of the question and the format of the chosen table in order to decide whether transferring is necessary, drawing graphs, finding values in the table, and doing some numerical calculations, etc. Some of the actions involve the cycling, integrating and generating tasks (Kirsch *et al*, 1998).

Some steps in the four stages could occur simultaneously, some steps could be in a linear time sequence and some steps might be revisited several times, so there is a need to find some way other than narratives to represent these steps and allow readers to have a clearer view of them. Harry *et al.* (2004) argue that visual display can be beneficial to researchers and instructive to readers. Therefore, in order to explain and analyse the participants' solving situations, I will express the participants' solving steps in a visual form.

With some adjustment, such a framework also applies to the questions of using the table of t distribution, which are Q3.1-Q3.2, and other questions of using the tables of normal distribution, which are Q1.5-Q1.6, Q15 in the first interview, Q1.a-Q1.d and Q6-Q7 in the second interview and Q9-Q10 in the third interview. Harry *et al.* (2004) point out that any visual representation of a complicated cognitive process is a vast simplification of the way that researchers actually arrive at interpretations. My visual presentation of the solving steps moves mainly from the description of the question on the left-hand side toward the right-hand side, through the given information and interpretation of the information, to the series of actions taken to deal with the interpreted information with the students' preferred/predetermined table. The actions moving downward represent the participants' major strategy and operation to obtain the z score for searching and its corresponding probability and possible further steps to get the final result. Therefore, in terms of the instruction which was received by the participants, a standard (but not necessarily the only correct) solving procedure of Q1.1-Q1.4 in basic formats can be displayed in Figure 5.2.

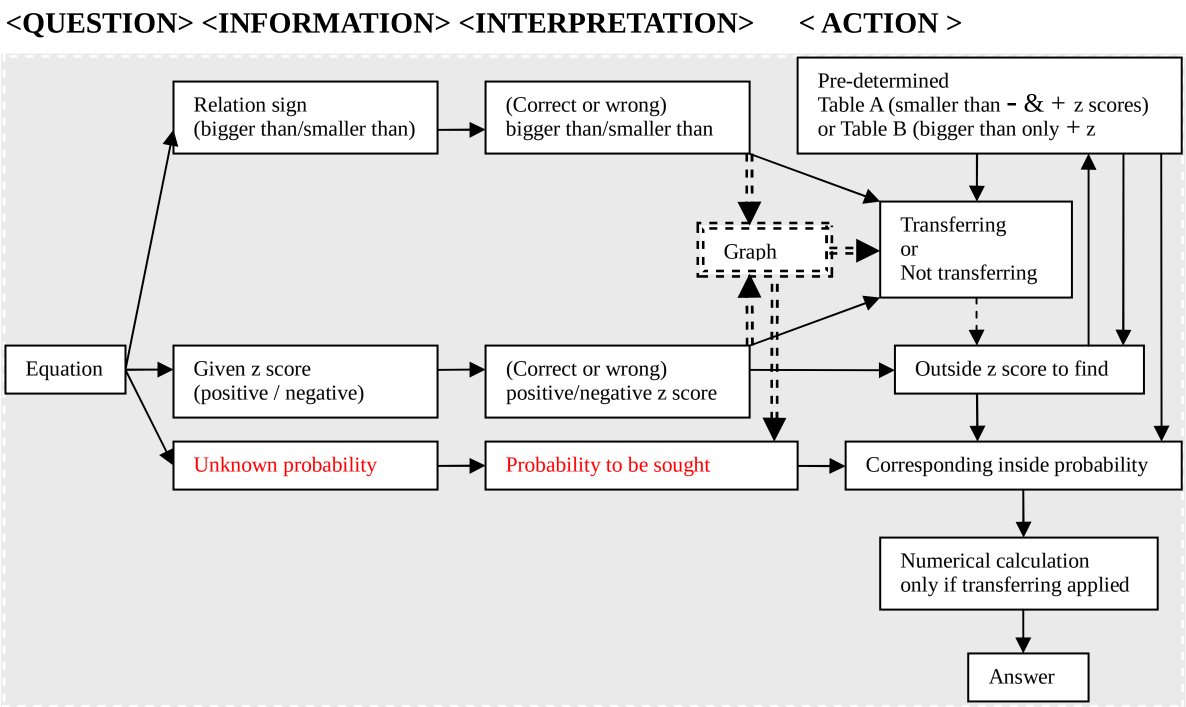


Figure 5.2. Solving processes of using the tables of normal distribution to find the probability according to the given z score.

**The dotted line means that transferring may or may not influence the value which should be found in the chosen table. This dotted line can be ignored if transferring is not applied.*

***The 'graph' in broken paralleled lines and the three double-dotted lines mean that graphs may or may not be used in the solving processes. They can be ignored if graphs are not applied in the solution.*

****The grey background means that this is one of the generalised solving processes of such questions.*

The answers for Q1.5 and Q1.6 cannot be found in the tables of normal distribution directly, and the original equation has to be transferred to the difference of two basic forms (see Figure 5.3.).

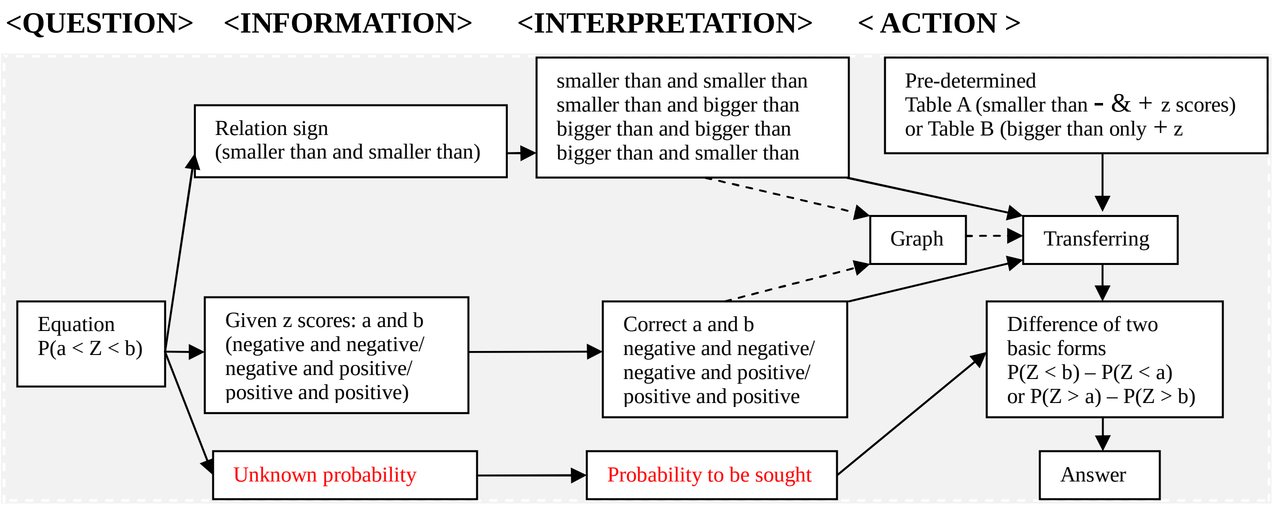


Figure 5.3. Solving processes of using tables of normal distribution to find the probability of Z between two z scores.

**The graph in dotted frame and its three dotted lines mean that graphs may or may not be used in the solving processes. They can be ignored if graphs are not applied in the solution.*

***The grey background means that this is the generalised solving process.*

However, there is another direction, ‘inside-out’: to find the given probability inside the tables of normal distribution and then correspond it to outside z score. This is expressed by Figure 5.4.

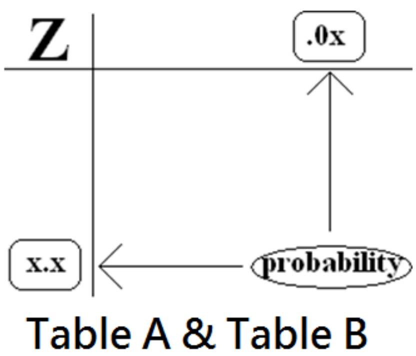


Figure 5.4. Graphical representation of using the tables of normal distribution in the direction of ‘inside-out’.

According to the teacher of the statistics course, the ideas of inside-out were taught in the sixth week of the second semester when he introduced the three commonly used z scores of confidence interval, but they were not reviewed or tested in the classes or examinations during the data collection period. Thus, such ideas were tested in Q15 in the first interview. Moreover, inside-out questions and outside-in questions were displayed together in Q1.a-Q1.d in the second interview to explore whether the participating students distinguished the question format and how they realised the meaning and relationship between the elements of the tables of normal distribution. I will discuss the solving process of such questions in Chapter 6.

5.3 Students' application of the table of t distribution

Another type of question involving 'outside-in' concerns the use of the table of t distribution. Although some participants used Table B and transferring principles well to solve Q1.1-Q1.6, it did not mean that they realised the meanings of the elements of Table B. Their interpretation and impression of elements inside or outside the tables of normal distribution can be highlighted by comparing their interpretation and application of the table of t distribution. T distribution was rarely used in the participants' classes, and it was not tested in their examinations. Therefore, students were not familiar with t distribution and the table of t distribution. In Q3.1-Q3.2, participants were asked to find a value according to symbolic representations, but they might not interpret the answer as the inside t score which could be found according to the given outside v (degree of freedom) and probability. The outside-in approach can be expressed by Figure 5.5.

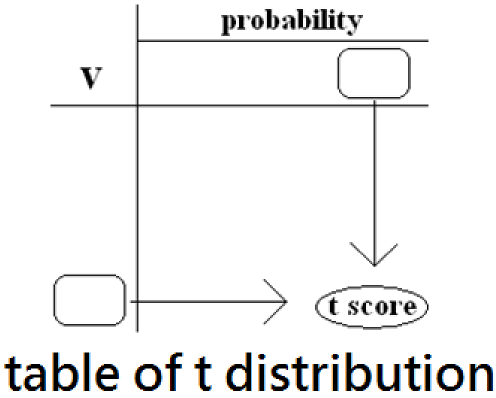


Figure 5.5. Graphical representation of using the table of t distribution in the direction of 'outside-in'.

There is only one kind of table of t distribution with a right-tail shadow area ($P[T > t_{\alpha}(v)] = \alpha$), so the issue is not to test the participants' preference, but to observe how they interpret the elements inside and outside the table and how they transfer the equation if the given information cannot be directly found in the table. In the participants' class, only outside-in questions were taught and practised, so no participant used the table of t distribution in the direction of inside-out in the interview. Typically, the solving process of looking for t score inside the table of t distribution according to given probability and given v can be expressed by Figure 5.6. I will discuss the participants' situations and difficulties of using the table of t distribution in Chapter 7.

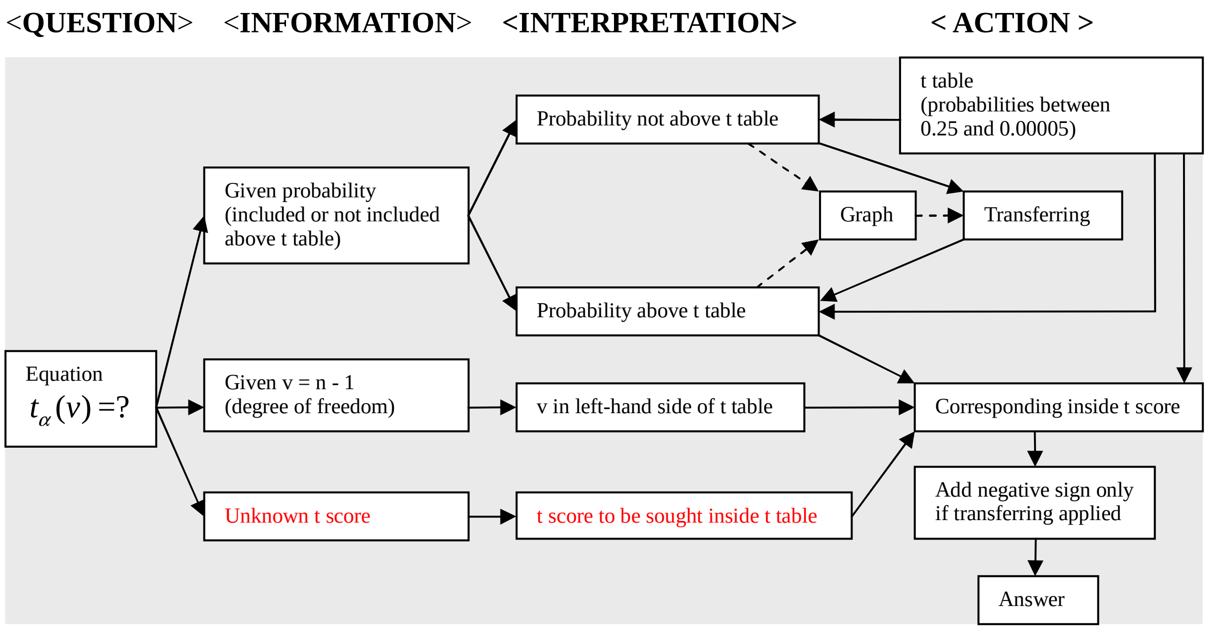


Figure 5.6. Solving processes of using t table to find t score according to given probability and degree of freedom.

**The dotted lines mean that participants may not take advantage of graphs when the given probability is not included above the t table and transferring is needed, and that they may still use graphs when the given probability is included above the t table and transferring is not needed.*

***The grey background means that this is the generalised solving process.*

5.4 Transferring principles of the tables of distribution

The participants solved Q1.1-Q1.6 by using the attached tables of normal distribution (i.e. Table A or/and Table B) and the memorised transferring principles, and a few of them used graphical methods to help themselves think of how the questions should be transferred. The transferring principles were first taught in the class by the teacher prior to my data collection, and he then asked the students to memorise them. The transferring principles of Table B and Table A were reviewed by the teacher in the eighth week when preparing for the mid-term examination.

5.4.1 Transferring principles of Table B

The transferring principles for the users of Table B are as follows:

Assume that 'a' is a positive value,

P(Z > a): use Table B directly;

P(Z < a) = 1 – P(Z > a);

P(Z < -a) = P(Z > a);

P(Z > -a) = 1 – P(Z > a); and

assume 'b' < 'c', P(b < Z < c) = P(Z > b) – P(Z > c).

The transferring principles are based on the idea of area. On the same page that Table B appeared in the textbook, there is a graph of a Z curve with a shadow area in the right-tail (shown in Figure 5.7.). It is noted in the graph that the right-tail shadow is bordered at z_α, and the area of the shadow area is α. An equation explaining the relationship between z_α and α is shown next to the graph as $P(Z > z_\alpha) = \alpha$, which means that the probability of Z being bigger than z_α equals α. The probability is also equal to the area of the shadow in the right-tail. The areas can be obtained by complicated calculation of integral, but it is not necessary to calculate them since four hundred probability values corresponding to z scores are included in Table B.

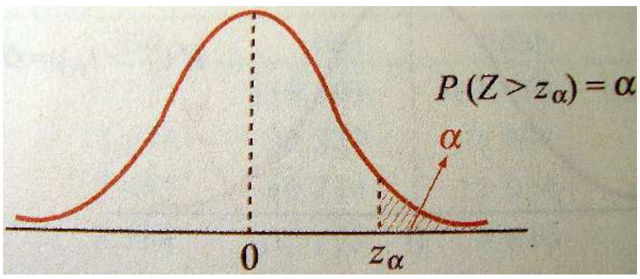


Figure 5.7. Graph of Table B attached in the textbook.

In the first teacher interview, the teacher reported that when he first introduced the transferring principles in class, he explained every principle by drawing graphs with the Z curve above a number line with a shadow region to emphasise the idea of area.

The first principle is in the basic format of Table B, so directly using Table B is adequate to solve questions in such a format. Questions in another format need to be transferred to such a format before using Table B. Thus, being able to recognise and distinguish the format of questions is important in solving the questions.

The second principle involves the idea of 'complementary'. That is, the total area below the Z curve equals 1, and since the area in the right-hand side is P(Z > a), the area of the other part in the left-hand side will be 1- P(Z > a). When the students were asked to find P(Z < a), they should first find P(Z > a), and subtract the found probability from 1. The graphical expression of this transferring principle is shown in Figure 5.8.

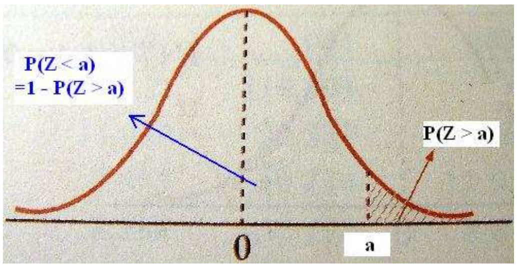


Figure 5.8 The graphical explanation of P(Z < a) = 1 – P(Z > a)

The third transferring principle involves another idea of area: 'symmetry'. Both sides of the Z curve are symmetric at z_α = 0, so the area of each side below the Z curve is 0.5. That is P(Z < 0) = P(Z > 0) = 0.5. Also, P(Z < -0.01) = P(Z > 0.01), P(Z < -0.02) = P(Z > 0.02),…P(Z < -3.99) = P(Z > 3.99), and they can be generalised as P(Z < -a) = P(Z > a). This principle is expressed in Figure 5.9.

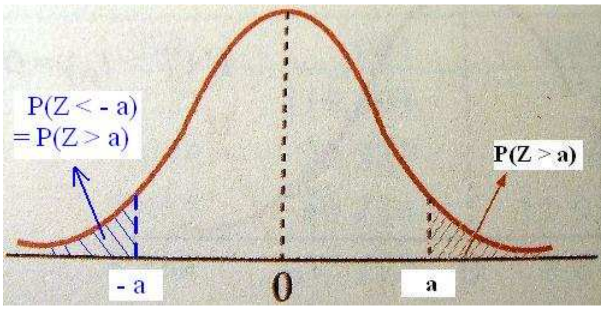


Figure 5.9. The graphical explanation of P(Z < -a) = P(Z > a).

The fourth principle requires both ideas of complementary and symmetry, and it is more complex than the previous ones. The format of P(Z > -a) can be transferred in two ways. First, it can be transferred to 1 – P(Z < -a) by applying the idea of complementary, and then transferred again to 1 – P(Z > a) according to the idea of symmetry. Transferring P(Z > -a) in this way can be visually expressed by Figure 5.10.

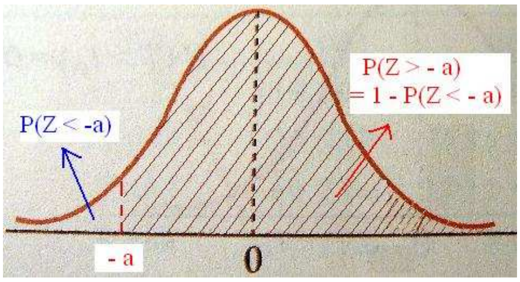


Figure 5.10. Graphical explanation of P(Z > -a) = 1 – P(Z < -a).

The other strategy is to transfer P(Z > -a) to P(Z < a) by applying the idea of symmetry first, and then use the idea of complementary to obtain 1 – P(Z > a) (see Figure 5.11.).

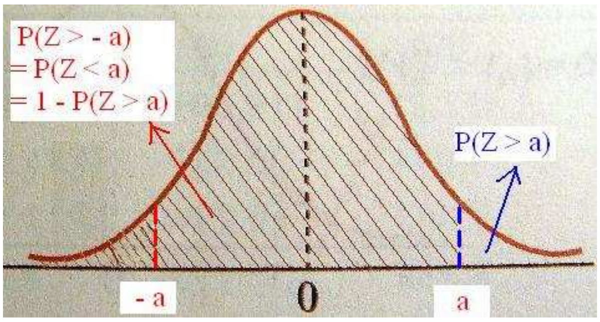


Figure 5.11. Graphical explanation of P(Z > -a) = P(Z < a) = 1- P(Z > a).

The fifth principle is only used in the advanced questions of betweenness. It should be first transferred to two basic forms according to the idea of area. The transferred equations will be transferred again only if one or both of them include(s) a negative z score. Assuming b is negative and c is positive, how P(b < Z < c) is transferred to P(Z > b) – P(Z > c) can be explained by Figure 5.12.

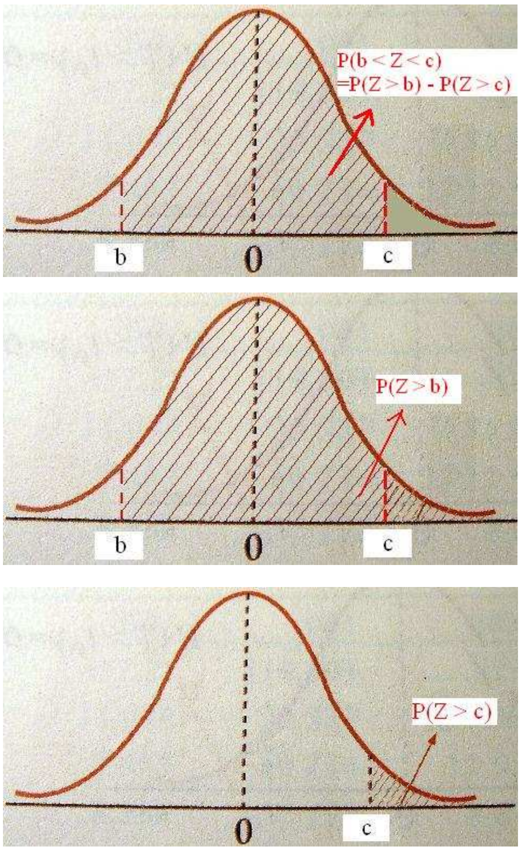


Figure 5.12. Graphical explanation of P(b < Z < c) = P(Z > b) – P(Z > c).

In Figure 5.12., we can see that the area of P(Z > b) is the sum of P(b < Z < c) and P(Z > c); that is,

P(Z > b) = P(b < Z < c) + P(Z > c).

Therefore, we can get

$P(b < Z < c) = P(Z > b) - P(Z > c)$

If both b and c are positive, the two transferred formats $P(Z > b)$ and $P(Z > c)$ can be found in Table B. If b is negative or both b and c are negative, the transferred format(s) will have to be transferred again, and this might cause students to make mistakes in calculation and symbolic manipulation. No matter if b is negative and c is positive, both b and c are negative or both of them are positive, the graphs will vary according to the given b and c, but the area of Z between b and c will still be the difference of the area of Z bigger than b and the area of Z bigger than c. However, many students relied totally on the memorised principles without understanding the origins and refused to use graphs when transferring, so they might not have been able to recall the correct formula when transferring was necessary. Q1.5 and Q1.6 will examine how the participants dealt with such questions.

5.4.2 Transferring principles of Table A

Although Table A was not often used in the class and was only used by S4 in the first interview, the teacher also introduced the transferring principles of Table A as follows.

Assume that ‘a’ and ‘b’ can be negative, positive or zero, and $a < b$

$P(Z < a)$: use Table A directly;

$P(Z > a) = 1 - P(Z < a)$ according to the idea of complementary

or $= P(Z < -a)$ according to the idea of symmetry; and

$P(a < Z < b) = P(Z < b) - P(Z < a)$.

In the first page of Table A in the textbook, there is also a graph of Z curve with a shadow area in the left-tail (shown in Figure 5.13.). It is noted in the graph that the left-tail shadow is bordered at z and the area of the shadow area is A. An equation explaining the relationship between z and A is shown next to the graph as $P(Z < z) = A$, which means that the probability of Z being smaller than z equals A. The probability also equals the area of the shadow in the left-tail. The areas also can be obtained by calculation of integral, but it is not necessary to calculate them since eight hundred probabilities are offered by Table A.

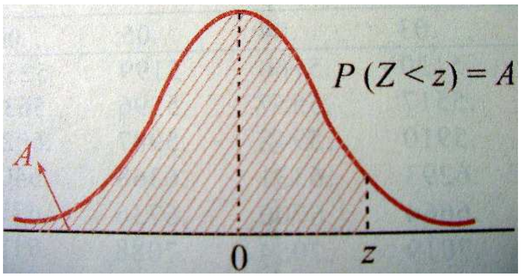


Figure 5.13. Graph of Table A attached in the textbook.

Unfortunately two factors might interfere with the participants' understanding of the attached graph in their textbook. First, the graph with a positive z score and a large shadow area was presented above the first half of Table A, which only includes negative z scores. Also, the second half of Table A, which includes positive z scores, was on the other page. That is, readers of this textbook could not see the entire Table A at the same time (see Figure 5.14. and Figure 5.15.). Thus, some readers did not notice the existence of the second half of Table A (positive z scores), and it is difficult for students to connect the graph with positive z scores and the table without positive z scores.

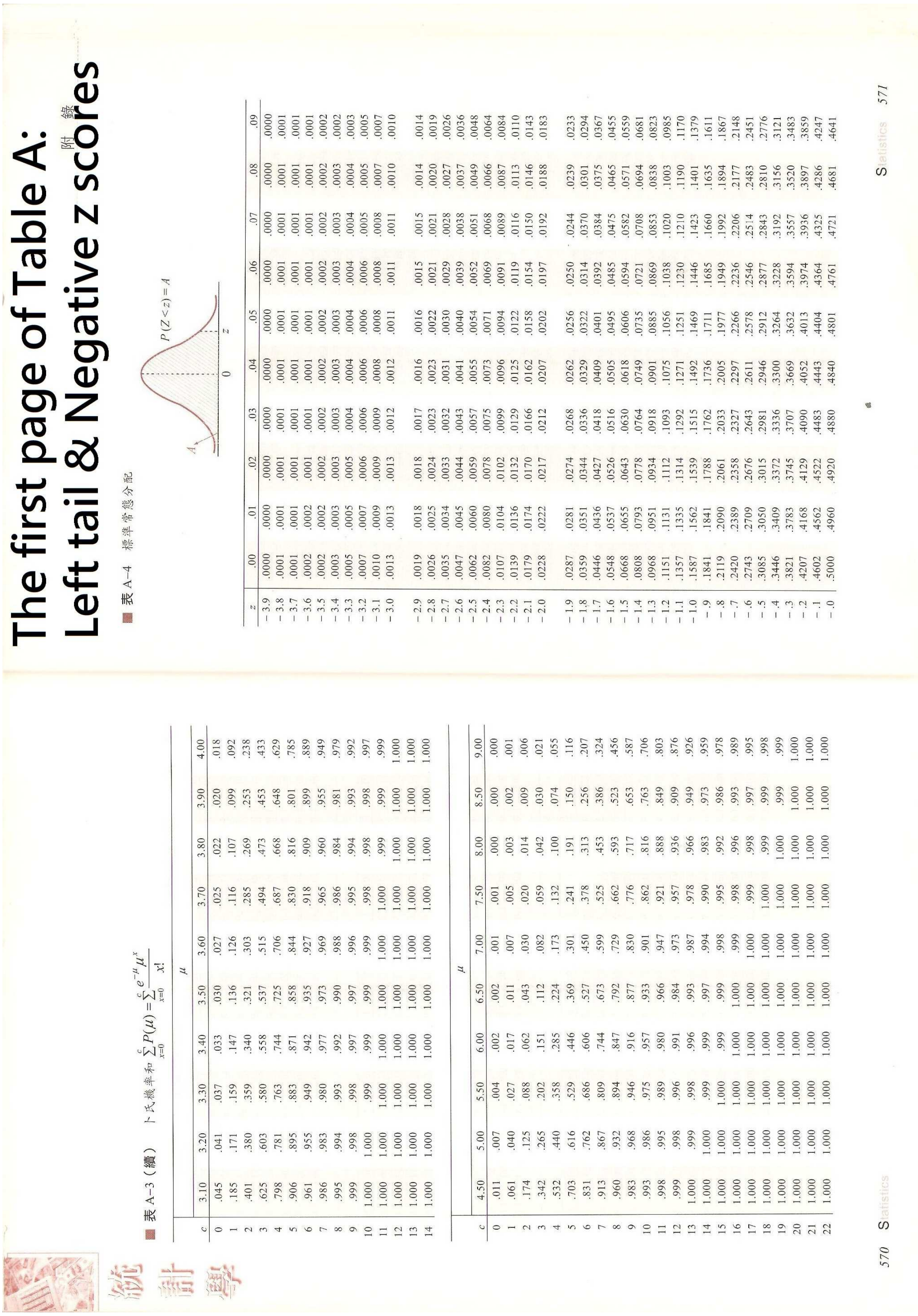
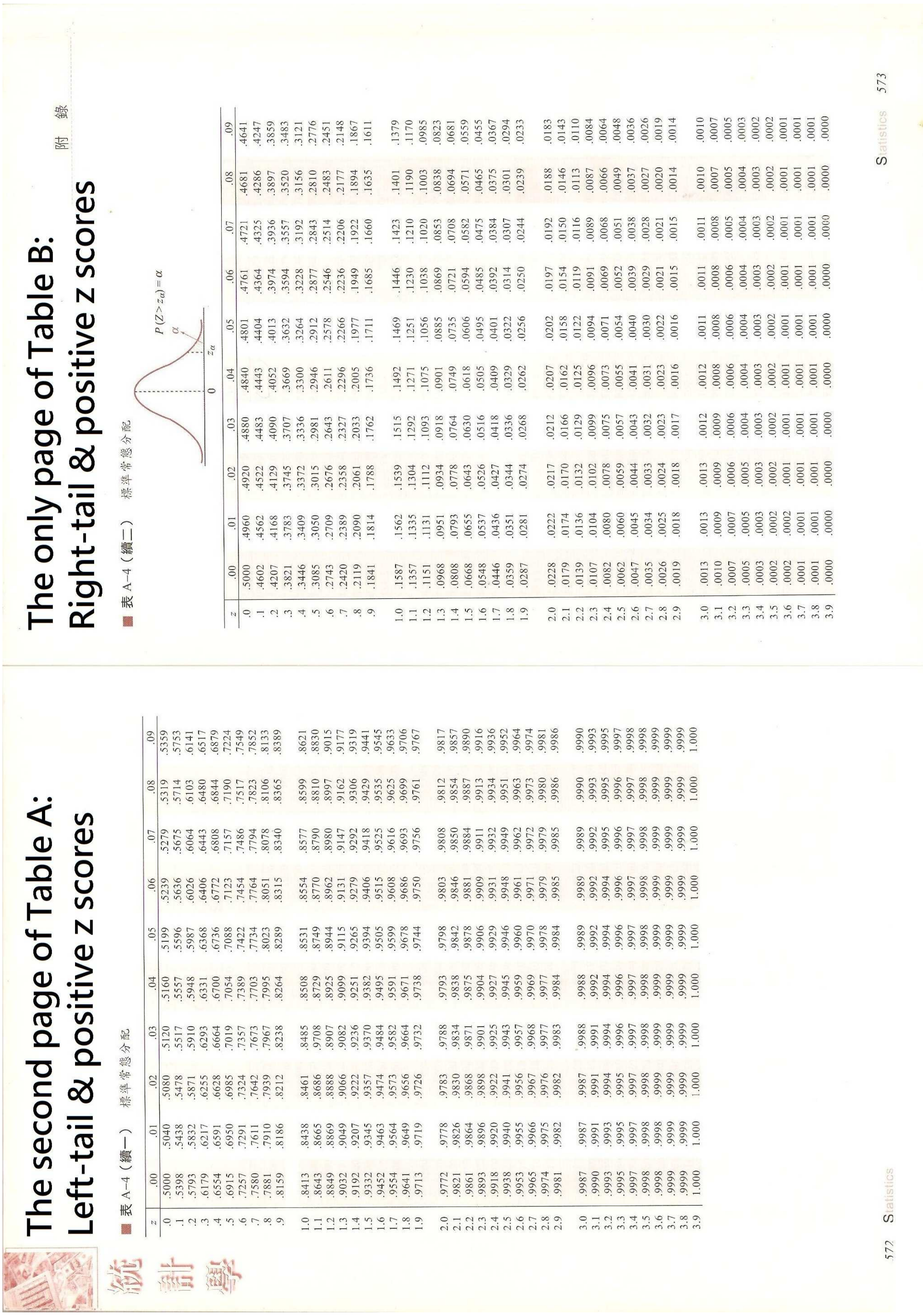


Figure 5.14. The table of Poisson distribution and the first page of Table A (Kuo and Shi, 2006).



The second page of Table A: Left-tail & positive z scores

■ 表 A–4（續一） 標準常態分配

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.9708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.8	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

The only page of Table B: Right-tail & positive z scores

■ 表 A–4（續二） 標準常態分配

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1304	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Figure 5.15. The second page of Table A and the whole Table B (Kuo and Shi, 2006.)

Second, the symbols in the graph of Table A are different from the symbols in the graph of Table B. Therefore, students who are used to the graph of Table B may not easily understand the graph of Table A. The teacher of the course explained that he found the problem of the attached graph and drew the ‘correct’ graph when he introduced the transferring principles of Table A. However, he rarely used Table A, so he did not pay much attention to whether and how the attached graph influenced the users of Table A.

The first principle is in the basic format of Table A, so directly using Table A is sufficient for users to solve questions in such a format. Because Table A includes both negative z scores (on the first page) and positive z scores (on the following page), questions in the format of P(Z < a negative value or zero) can be found on the first page of Table A (see Figure 5.16.), whereas questions in the form of P (Z < a positive value or zero) can be found on the second page (see Figure 5.17.). Questions in other formats also need to be transferred to such a format before using Table A.

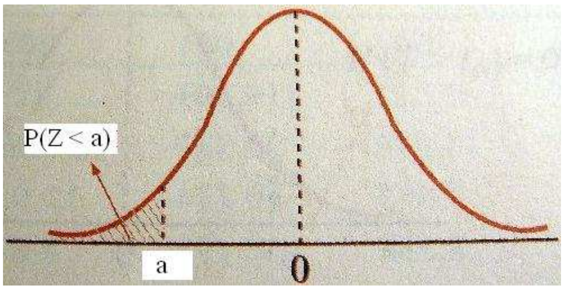


Figure 5.16. The graphical explanation of P(Z < a) if a $\leq$ 0.

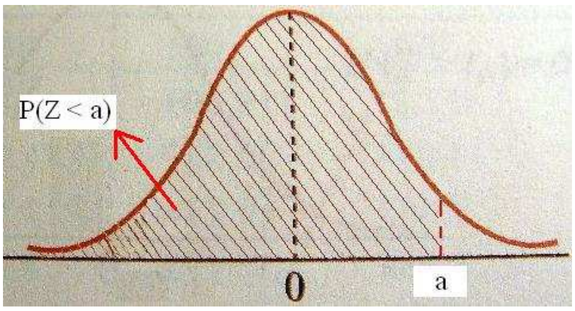


Figure 5.17. The graphical explanation of P(Z < a) if a $\geq$ 0.

The second principle deals with questions of P(Z > a value), and there are two ways to transfer the original format: complementary or symmetry. The ideas of complementary and symmetry have been introduced in the second and the third transferring principles of Table B. Like the first principle, the graph of the second transferring principle of Table A is also influenced by whether the given value is positive or negative. If ‘a’ is positive or zero, P(Z > a) can be transferred to 1 - P(Z < a) by the idea of complementary, as is explained in Figure 5.18.

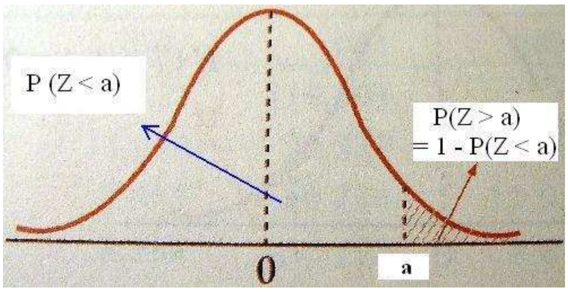


Figure 5.18. The graphical explanation of P(Z > a) = 1 – P(Z < a) if a $\geq$ 0.

P(Z > a) also can be transferred to P(Z < -a) according to the idea of symmetry, and this is graphically expressed in Figure 5.19.

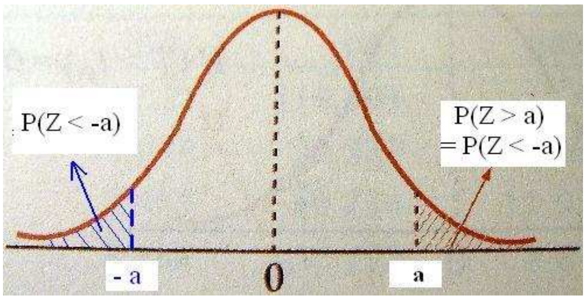


Figure 5.19. The graphical explanation of P(Z > a) = P(Z < -a) if a $\geq$ 0.

If ‘a’ is negative or zero, P(Z > a) can be transferred to 1 - P(Z < a) by the idea of complementary, as explained in Figure 5.20.

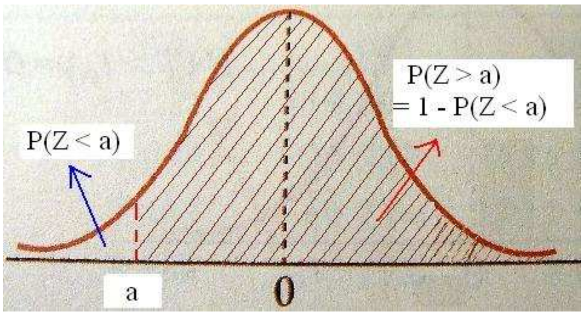
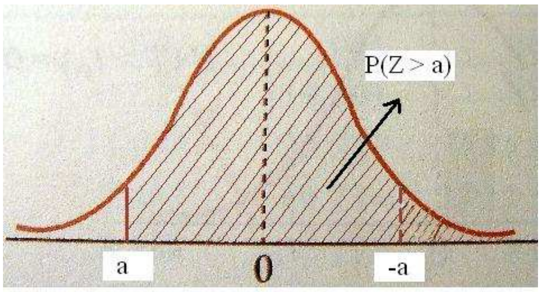


Figure 5.20. The graphical explanation of P(Z > a) = 1 – P(Z < a) if a ≤ 0.

P(Z > a) also can be transferred to P(Z < -a) according to the idea of symmetry, as is graphically expressed in Figure 5.21. However, some students may still see ‘-a’ as negative because they believe that “a minus sign means a negative number” (Schechter, 2009, p.6).



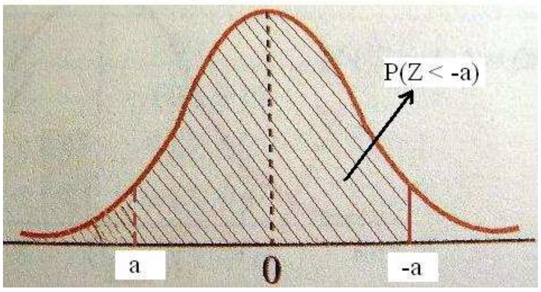


Figure 5.21. The graphical explanation of P(Z > a) = P(Z < -a) if a ≤ 0.

The third transferring principle of Table A can be explained as in Figure 5.22. Unlike the fifth transferring principles of Table B, only one transferring process is required for Table A users since either negative z scores or positive ones can be found in Table A.

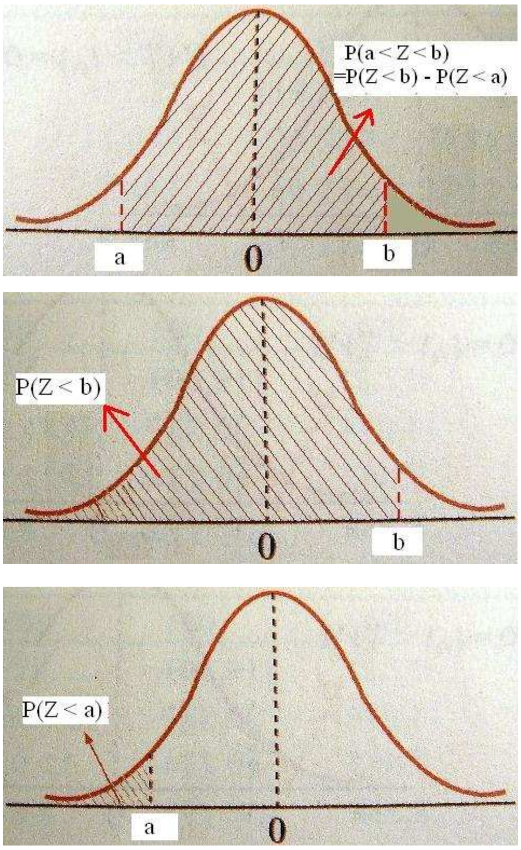


Figure 5.22. The graphical explanation of $P(a< Z < b) = P(Z < b) – P(Z < a)$.

5.5 Comparison of Table A and Table B

Both Table A and Table B could be used by the participants (except for S1 and S2, who were only given Table B) to solve Q1.1-Q1.6 in the first interview, but all of the participants showed their preference for a particular table. It is worth mentioning that S1 preferred Table A but only used Table B because Table A had not been given to him in the first interview and he had not asked for it. He explained that he did not ask for Table A because he thought the interview was testing how he used Table B. Only S4 used Table A, while the other nine participants used Table B to solve Q1.1-Q1.6. During the first interview, very few participants considered using the other table even if the answers of some questions could be directly found in it. Instead, they relied on the memorised transferring principles to transfer the question to a suitable format before using the pre-determined table.

The format of Table A [$P(Z < z_\alpha) = \alpha$] represents the probability of random variable Z being smaller than a negative or positive z score, whereas Table B [$P(Z > z_\alpha) = \alpha$] represents the probability of Z being bigger than a positive z score. Therefore, transferring steps became inevitable for users of Table B when the question involved negative z scores.

Some students may transfer the equations only by symbolic manipulation, and some others may transfer the equations according to graphs. No matter whether transferring was correctly applied, mistakenly applied or not applied at all by an individual, the next stage was to use the table. The z score for a participant to look for outside the table was not necessarily the given z score in the question, but was determined by the combination of his/her interpretation of the relation sign and z score and the chosen table.

When the outside z score was found, the participant could correspond it to a probability inside the table. The z score was separated into two parts: 'something-point-something' (i.e. the integer and the first digit in the decimal part) on the left-hand side of the table and 'point-zero-something' (i.e. the second digit of the decimal part) on the top of the table. Then the corresponding probability was located at the intersection of the row of something-point-something and the column of point-zero-something. For instance, the intersection of the row of '1.9' on the left-hand side and the column of '.06' on the top of Table B is not 1.96 but .0250, which means that the probability of Z being bigger than 1.96 is 0.025, whereas the intersection of the row of '1.9' on the left-hand side and the column of '.06' on the top of Table A is .9750, which means that the probability of Z being smaller than 1.96 is 0.975 (see Figure 5.23.).

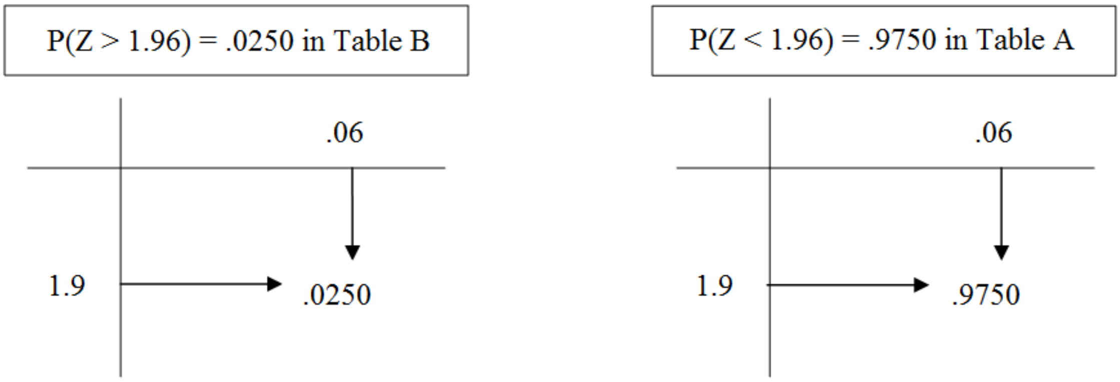


Figure 5.23. Comparison of the intersection of the row of 1.9 and the column of .06 in Table B and in Table A respectively.

If transferring was not applied, the found probability would be the answer; otherwise, the found value had to be put back to the applied transferring principles in order to get the final answer. For example, the answer to Q1.4 was the found '0.0202' inside Table B; however, the answer to Q1.3 was not simply the found '0.001' inside Table B but rather '1 – 0.001', since transferring was applied in order to shift the question to the format of Table B.

There are several differences between Table A and Table B. First, since Table B only includes the probability of Z being bigger than positive z scores or zero (from 0.00 to 3.99) while Table A covers the probability of Z being smaller than not only negative z scores and zero but also positive z scores (from -3.99 to 3.99), the scope of Table A is two times that of Table B. S1 explained that he preferred Table A because of its bigger range (i.e. more probabilities).

Second, since Table A includes more data than Table B, Table A enables its users to find more values in it 'directly'. For example, in the four basic formats of question Q1.1-Q1.4, the answers to Q1.1 and Q1.2 can be found directly in Table A whereas only the answer to Q1.4 can be found directly in Table B. For the questions whose answers cannot be found directly in Table A or Table B, transferring processes have to be undertaken. Therefore, the users of Table A have to transfer Q1.3 and Q1.4 before using Table A, whereas the users of Table B need to transfer Q1.1, Q1.2 and Q1.3 first.

Transferring steps are unavoidable before using Table A or Table B in the advanced questions Q1.5 and Q1.6, but Table B users may need more transferring than Table A users. The transferring principles and theoretical grounds for solving Q1.1-Q1.6 are listed in Table 5.1.

	Questions	**Principle and theoretical basis for Table A users**	**Principle and theoretical basis for Table B users**
Q1.1	P(Z < –0.03)	N/A (answer directly found).	= P(Z > 0.03): involves symmetry.
Q1.2	P(Z < 1.22)	N/A (answer directly found).	= 1 – P(Z > 1.22): involves complementary.
Q1.3	P(Z > –3.1)	=1 – P(Z < –3.1): involves complementary or = P(Z < 3.1): involves symmetry.	=1 – P(Z < –3.1) = 1 – P(Z > 3.1): involves complementary and symmetry or = P(Z < 3.1) = 1 – P(Z > 3.1): involves symmetry and complementary.
Q1.4	P(Z > 2.05)	= 1 – P(Z < 2.05): involves complementary or = P(Z < –2.05): involves symmetry.	N/A (answer directly found).
Q1.5	P(–1.4 < Z < 2.32)	= P(Z < 2.32) – P(Z < –1.4): involves a middle area equal to the difference of two left-tailed areas, get two basic forms of Q1.2 and Q1.1 and then simply use subtraction.	= P(Z > –1.4) – P(Z > 2.32): involves a middle area equal to the difference of two right-tailed areas, get two basic forms of Q1.3 and Q1.4, transfer P(Z > –1.4) to 1 – P(Z > 1.4) with the

			principles of Q1.3 and then use subtraction.
Q1.6	P(–2.41 < Z < –0.03)	= P(Z < –0.03) – P(Z < –2.41): involves a middle area equal to the difference of two left-tailed areas, get two basic forms of Q1.1 and then simply use subtraction.	= P(Z > –2.41) – P(Z > –0.03): involves a middle area equal to the difference of two right-tailed areas, get two basic forms of Q1.3, transfer P(Z > –2.41) to 1 – P(Z > 2.41) and transfer P(Z > –0.03) to 1 – P(Z > 0.03) with the principles of Q1.3 and then use subtraction. Beware of the change of minus sign.

Table 5.1. Transferring principles and theoretical backgrounds for Table A users and Table B users.

By viewing this table, we can see that in solving the four basic questions, the Table A users can more often directly find the answers (2/4) than the Table B users (1/4). In other words, the Table B users (3/4) have to transfer before using the table more often than the Table A users (2/4). The biggest difference between the Table A users and the Table B users can be seen in the solving process of Q1.5 and Q1.6. When solving Q1.5, the Table A users only have to transfer the question to two basic forms (like those of Q1.2 and Q1.1 respectively, both of which can be found in Table A) and then simply use subtraction to get the answer; Table B users have to transfer the question to two basic forms (like those of Q1.3 and Q1.4, respectively), but only the answer of the latter form (of Q1.4) can be directly found in Table B, so an extra transferring process (like that in Q1.3) is needed before using Table B and applying subtraction. Solving Q1.6 is even simpler for Table A users because Q1.6 will be transferred to two basic forms of Q1.1. However, this question seems to be the most challenging for Table B users. The first step of solving Q1.6 is to transfer it to two basic forms of Q 1.3, and both of them need to be transferred again with the principle used in Q1.3 (i.e. this question requires three transferring steps). The change of the minus sign resulted in a problem for a participant to get

P(Z > –2.41) – P(Z > –0.03) = 1 – P(Z > 2.41) – 1 – P(Z > 0.03).

Whoever was able to switch to the corresponding table according to the type of question could use the transferring principles in fewer situations. For instance, he/she may directly find the answer of Q1.1 and Q1.2 in Table A and the answer of Q1.4 in Table B, so he/she only has to transfer Q1.3. I will observe which table was used by the participants and explore the reasons and influences in the clinical interviews.

Although there were fewer transferring principles for the Table A users than for the Table B users, and although Table A enables users to find values in it directly more often than Table B, the teacher normally used Table B, so Table A was not used by many students in the first interview.

Furthermore, the participants were not familiar with the t distribution because the teacher did not let students use the table in classes and examinations. Even if the idea of t distribution was needed in the final examination, the corresponding t scores and probabilities were offered by the teacher in the question sheet as follows:

$$Z_{0.975} = 1.96, Z_{0.99} = 2.325, Z_{0.995} = 2.575,$$
$$t_{0.01}(10) = 2.764, t_{0.01}(11) = 2.718, t_{0.1}(11) = 1.363$$
$$S_p{}^2 = \frac{(n_1 - 1)S_1{}^2 + (n_2 - 1)S_2{}^2}{n_1 + n_2 - 2}$$

Therefore, the transferring principles of Table B will be the major part of my discussion.

5.6 Summary

In this chapter, I first introduced the content of the course, such as the questions in practice, mid-term and final examinations, the textbook used in the course, and the concepts taught in the classes. Then I discussed the students' usual way of using the tables of normal distribution and the table of t distribution and developed three frameworks of solving steps in order to explain the process of using the tables of distribution to solve three kinds of questions. These solving steps enabled me to observe each step of the participants in their solutions. I also discussed and compared the transferring principles of using Table B and Table A and compared the two tables in the participants' textbook. In the next four chapters, I will use the framework of solving steps to analyse how the participants used the tables of distribution and transferring principles to solve the exercise questions in the clinical interview.

CHAPTER 6

DESIGN OF THE CLINICAL INTERVIEWS

In this chapter, I will explain the process of designing questions for the teacher interviews and student interviews. I will discuss some findings from the class observations and the materials analysis, explain how these findings relate to the study and then link the findings to the questions for interviews. These interview questions and the participants' responses will provide evidence to explore the research questions of the study. For the questions and transcripts of the interviews, see Appendix 6.1. - Appendix 6.7. and Appendix 6.11. - Appendix 6.19.

6.1 Teacher interview I (week 9)

The first teacher interview was conducted in the ninth week of the second semester. It began with a few questions, concerning issues such as the composition of the students in the statistics course, the teacher's experience of teaching statistics and the choice of textbook. These questions aimed to garner a better understanding of the background of the class.

6.1.1 Part I of the first teacher interview

The first part of the interview explored the teacher's attitude towards the use of visualisation in his statistics learning and teaching. I asked the teacher to compare the advantages of visual form and non-visual form (symbolic manipulation). The teacher is a graph user, and he believes that using graphs is important for understanding the tables of normal distribution, but drawing graphs is not compulsory in his class.

According to the class observation in the eighth week, when the teacher announced that most of the practice questions (see Appendix 4.5.) would be tested again in the mid-term examination (see Appendix 4.7.), it seems possible that some students would simply memorise the answers to the questions without understanding them fully. I therefore asked the teacher why he offered the questions and solutions to the students before the mid-term examination. The teacher explained that he was worried about the students' performance and did not want them to fail, so he let the students practise to get familiar with the questions.

I examined the preparatory paper and noticed that the paper did not include the t distribution, F distribution or Chi-square distribution in the mid-term examination. The teacher explained that these were too difficult for the students.

6.1.2 Part II of the first teacher interview

In the second part of the interview, I asked why the teacher used the table of normal distribution with right-tail (i.e. Table B) rather than the table of normal distribution with left-tail (i.e. Table A) since both tables were provided in the textbook. He preferred Table B because his teachers had used Table B and he was used to it. I also asked the teacher to compare the advantages and disadvantages of the two tables and the other table of normal distribution with a shadow area between zero and a positive z score (see Figure 6.1.). Such a table was not included in the textbook, but was used in many other textbooks such as Bluman (2008) and Liu *et al.* (2005).

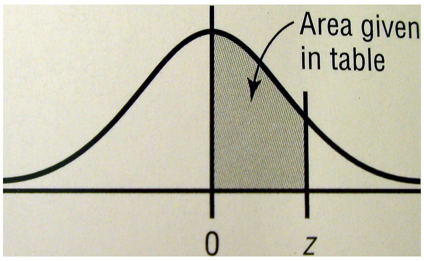


Figure 6.1. The graph of the table of standard normal distribution with a shadow area between zero and a positive z score.

6.1.3 Part III of the first teacher interview

In the third part of the interview, I, according to the class observation in the eighth week, asked the teacher why he wanted the students to memorise the principles of transferring rather than to transfer the formats with graphs, and why he asked the students to memorise the z scores of confidence interval rather than to look for them in the table of normal distribution. The teacher explained that he wanted the students to memorise the three frequently used z scores for convenience and that this might be beneficial for their scores.

I analysed the question sheets in the preparatory paper and in the mid-term test and found that the step of standardisation was used nine times, so there seemed to be a necessity to explore students' understanding of standardisation. In addition, I asked the teacher to explain how he taught the idea of standardisation in class. The teacher said that he explained standardisation in class as "a step to find probability in the tables of normal distribution" (23rd April 09), and he recognised that students might not understand the real meaning of standardisation. Thus, I decided to explore the students' understanding of standardisation in the first student interview.

6.2 Student interview I (weeks 10 and 11)

The first student interviews were taken after the mid-term examination and after the first teacher interview. When I designed the questions of the first student interview, I took a few aspects into account. First, the teacher revealed the solutions of most questions to the students, and some students might simply memorise the answers to the questions instead of working to actually understand them. Therefore, I asked the participants whether they had solved the questions by themselves or had simply recalled the answers to the questions on the mid-term examination. More than half of the participants admitted that they had used some of the memorised solutions as the answers.

Second, I analysed the questions in the preparatory paper and mid-term examination, and I found that solving the questions required the students to recognise the conditions of the question in order to recall the formulas, and/or to find probabilities according to the given z scores or t scores after the step of standardisation. Since the answers had been revealed by the teacher, I decided to design new questions to collect data of how the participants used the tables of normal distribution and the table of t distribution and what problems they encountered. I also examined whether and how they used graphs of normal curve and t curve when using the tables. Additionally, seeing that standardisation was used nine times in the preparatory and mid-term paper, I chose to investigate the students' understanding of standardisation.

Third, I had observed the teacher ask the students to memorise transferring principles as formulas, and he also asked the students to remember z scores of confidence interval rather than to look for them in the table of normal distribution. In the first teacher interview, the teacher explained that he had asked the students to memorise three frequently used z scores for convenience, as this might be beneficial for their scores.

Therefore, I investigated how the students transferred the formats and how they used the tables of normal distribution.

In the first interview, there were four groups of questions which aimed to explore how students understood and used the tables of normal and t distributions, how students interpreted visualisation, which table of normal distribution they preferred and why, and how students interpreted standardisation and confidence interval. I will explain how I designed these questions in detail and discuss the participants' responses below.

6.2.1 Part I of the first student interview

The first student interviews began with a series of questions about using the tables of standard normal distribution to find the probabilities of standard normally distributed variable Z being bigger or smaller than a given value or between two given values. The skills and principles required to solve such questions had been taught in the first semester and frequently used afterwards. Garfield and Ben-Zvi (2008) suggest that distribution is one of the most important big ideas in statistics because it involves variability, sampling and inference.

In the teacher's mid-term questions, the table of normal distribution should be used in solving processes in ten sub-questions, but the explained solutions could be memorised. Therefore, I designed new questions to cover each kind of condition (such as Z being smaller than, bigger than or equal to a value, or between two values, and the value(s) being positive or negative) to explore the participating students' solving strategies. I gave the participants basic questions Q1.1-Q1.4 and advanced questions Q1.5-Q1.6 (see Appendix 6.3.) to explore 'which table they preferred', 'why they preferred it' and 'whether/how they transferred the question to a correct format for the chosen table'. The numbers used in the six questions were selected in the question designing process in order to diminish errors in inevitable calculation. The advanced question type of finding the probability of Z falling between two positive values (such as P ($1 < Z < 2$) =?) was omitted in order to save time for other interview questions. In addition, it was simpler and easier than Q1.5 and Q1.6.

Solving the six questions did not require complex calculation, but each participant had to decide whether the format of the question matched the table of normal distribution which he/she preferred to use according to the ‘relation signs’ (bigger than or smaller than) and the positive/negative ‘z scores’. If the two factors in the question and in the format of the predetermined table were the same, this table would be considered suitable to solve this question, and the table user could directly find the given z score outside the table and then find its corresponding probability inside the table as the answer; otherwise, the question should be transferred to the format of the chosen table.

According to the first teacher interview, the idea and the table of t distribution were also taught and used in the participants’ classes, but the teacher did not test these in the mid-term, so I included two questions of t distribution because some of its characteristics (only positive part and right-tailed) are similar to that of Table B. I paid attention to whether/how the participants distinguished between the tables. The first sub-question Q3.1 tested the basic principle of using the t table, whereas the second one Q3.2 involved symmetry and negative t value. I did not test tables of F and Chi-square distributions because they were not mentioned or used in the classes and examinations. However, I would plan to test them in the following interviews, once they had been used by the teacher in classes or on examinations.

The focus of the exercise questions in the first part was to see whether the students found the value(s) in the correct place(s) and whether the transferring principles were correctly applied when transferring processes were needed. The transferring principles were discussed in Chapter 5, and I will discuss how the students responded to Q1.1-Q1.6 and Q3.1-Q3.2 in Chapter 7 and Chapter 8 respectively.

6.2.2 Part II of the first student interview

In the first teacher interview, the teacher said that he believed it necessary to use graphs to understand the tables of normal distribution, and he had accordingly drawn graphs in previous classes. However, he was observed in the eighth week asking the students to memorise the transferring principles and z scores, and he did not appear concerned with whether the students learned the tables with graphs. Therefore, in the second part of the interview, I invited the participants to talk about their definition of visualisation and share their attitudes and experiences of using visualisation throughout the statistics course and in the specific chapters of statistical inference. I also asked them to review the previous chapters and discuss the advantages and necessities of visualisation. I will discuss the students' responses to these questions and link their attitude to their performance in Chapter 11.

6.2.3 Part III of the first student interview

The third part of the interview aimed to explore the students' preferences for tables of normal distribution and their reasons. There are three kinds of table of normal distribution commonly used by authors of statistics textbooks, but normally only one or two types was/were introduced in each textbook. For example, Wonnacott and Wonnacott (1990) use the table of right-tailed, and Bluman (2008) only adopts the table of the shadow area (α) between zero and a positive value. Table A (left-tailed with negative and positive z scores) and Table B (right-tailed with positive z scores) were included in the participants' textbook, and both of them were introduced and used by the teacher. Although the teacher often used Table B, he did not restrict the choice of table of normal distribution, so the students could use either table. The students' preferences and obstacles might be influenced by the condition and principle of different tables, so I asked the participants to explain their preferences and discuss the difficulties they had when using the table(s). This will be discussed in Chapter 7.

6.2.4 Part IV of the first student interview

In the preparatory paper, the idea of standardisation was tested in nine sub-questions, and the z scores of 95% and 90% confidence intervals were needed. However, the questions and solutions were revealed one week before the mid-term test, raising the possibility that the students might simply have memorised the numbers and answers without understanding their origin and using the tables of normal distribution to find them. The teacher also asked the students to remember the three z scores and their corresponding confidence intervals: 1.645 for 90%, 1.96 for 95% and 2.575 for 99%. Therefore, there was a necessity to examine how students interpreted the standardisation and confidence interval. The last part of the first student interview focused on this issue, so I asked the participants to find the frequently used z scores of confidence interval in the tables of normal distribution.

Wonnacott and Wonnacott (1990) explain the meaning of standardisation as follows:

> the Z value gives the number of standard deviations away from the mean (1990, p.131).

The teacher said that standardisation was used in his class as "a step to find probability in the tables of normal distribution" (the teacher, 23rd Apr 09), and he recognised that students might not understand the real meaning of standardisation. Therefore, I decided to explore students' understanding of standardisation, and I also examined how the participants find the z score of confidence interval. I will discuss the participants' responses to these questions in Chapter 9.

The audio- and video-recordings of the first student interviews were transcribed as soon as possible after they were taken. The transcripts were repeatedly revised in order to prevent including wrong information or missing useful information. The important information collected in the first student interview is listed, coded and categorised in Appendix 6.8. – Appendix 6.10., and this provided direction for further class observation and question design.

6.3 Student interview II (week 14 and 15)

The second student interviews were taken five weeks after the first interviews. I had noticed some phenomena in the first student interviews and the five-week class observations, so I designed questions in the second student interviews based on these findings. The second student interviews consisted of three parts.

6.3.1 Part I of the second student interview

The first part of the second interview was based on some clues explored in the first student interview, such as the un-noticed relation sign and the participants' difficulties of finding a probability which could not be seen in the table of t distribution. In the first interview, I asked the participants to find the probabilities in the tables of normal distribution according to the given relation signs and z scores to explore the students' choice of Table A or Table B and how they transferred the equations. I tested some types of the question again in the second interview to see whether the students changed, improved or forgot the solving strategies which they used in the first interview.

Sierpinska (1994) suggests that analysis of students' responses enables the researcher to "order the targeted aspects with respect to their difficulty" (1994, p. 119). When I reviewed the students' responses in the first interview, I wondered whether the participants were clear about the 'format' of the question, so I referred to the two questions at the end of my video (see Figure 1.1) to introduce new formats of questions and to test the participants' interpretation of the format of the equation and the elements in the tables of normal distribution, such as relation sign and location in the equation. I designed another series of questions (Q1.a - Q1.d in the second interview) in which I again manipulated the relations of 'bigger than' and 'smaller than' and added a reverse format of question in which participants should find the z score with a given probability value. These questions involved two reverse approaches of using the tables of normal distribution, which were 'looking for probability inside the table according to given (outside) z score' and 'looking for outside z score according to the given (inside) probability', and tested students' ability to solve the "inverse problems" (Dyrszlag, 1984, noted in Sierpinska, 1994, p.118).

Some participants might use the other method to find some anticipated wrong answers. Taking the first two questions in the second interview as an example, $P(Z > a) = 0.67$ (outside-in) was a new format for the participants, while $P(Z > 0.67) = b$ (inside-out) was the format tested in the first interview. The answer of 'a' cannot be directly found in Table A nor Table B, and this aimed to test whether the participants recognised and distinguished the formats of equation and how they transferred the format, such as using the idea of symmetry or drawing graphs.

For the sub-questions in this part, only the value '0.67' was used because the z score 0.67 could be directly 'found' in both Table A and Table B, and because probability 0.67 could also be 'found' inside Table A. The questions aimed to see if students distinguished or were confused by the question formats. With the same value 0.67, the sub-questions differed by two factors, which are the 'locations of the given value 0.67 and the unknown' and the 'relation sign'. By doing this, a new format of question, looking for z scores according to the given probabilities and the relation signs, could be used to see whether/how students distinguished the question format and how they interpreted the elements of the tables of normal distribution. I will discuss the participants' solutions to the questions in Chapter 7.

6.3.2 Part II of the second student interview

The second part of the second interview examined students' understanding of sampling distribution. The teacher paid much attention to the 'difference between two means' and the 'difference between two proportions' in class and examination. Therefore, in the first sub-question, I gave the participants $\overline{X} \sim N(20, 20^2)$ and $\overline{Y} \sim N(10, 10^2)$ and asked them to find the mean and variance of a normal distributed variable of difference between two means ($\overline{X} - \overline{Y}$). I allowed the participants to use the textbook in the interviews, and this question could be solved if they remembered the formula or found the formula in their textbook.

After collecting the participants' solutions to the first sub-question, I gave the participants $\overline{X} - \overline{Y} \sim N(10, 5^2)$, drawing the curve of normal distributed variables $\overline{X}$ as an example, and asked them to draw the curves of $\overline{Y}$ and $\overline{X} - \overline{Y}$. This question prompted participants to display their interpretation of mean and standard deviation of normally distributed variables. The participants' performances and explanations to their drawings will be analysed and discussed in Chapter 11.

I also asked the participants to find the 95% confidence interval of $\overline{X} - \overline{Y}$. This question could be solved by memorising/finding the formula and symbolic manipulation as well. Confidence interval was tested in the first interview, and it was tested again in the second interview because the formulas of confidence interval of difference between two means became much more complicated than that of a normal distributed variable. This question was followed by judging whether the two normally distributed variables were significantly different.

6.3.3 Part III of the second student interview

The third part of the second interview tested students' ability at solving hypothesis testing questions and explored their preference of approaches. I designed the questions according to my findings in the classroom observations, including some unexpected teaching problems.

In most situations, the statistics classes were observed to be typical of a Taiwanese classroom: the teacher stood speaking in front of the classroom, used and explained the content in the textbook, such as theories, real life examples and exemplar questions, and wrote on the blackboard, whereas the students sat and listened to the teacher with very few interactions with their classmates or the teacher. The teacher's lessons were always structured as follows:

Review the relevant prior knowledge, including ideas and formulas;

Introduce the new subjects by linking to the relevant contents which had been taught; and

Solve the exemplar questions in the textbook, by listing the conditions in symbolic ways and putting the numbers into the formulas.

The teacher used symbolic manipulation to solve each exemplar question he used in class. He was observed to use graphs in some types of questions, but he did not use graphs in all of these questions. Whether/how the used graphs influenced the teachers' solutions deserves to be examined. As displayed in Table 6.1., the teacher did not use any graph when teaching estimation in the observed classes, but he used graphs fifteen times in four consecutive weeks (12th, 19th, and 26th May and 2nd Jun 2009) when teaching hypothesis testing.

Week	Date	Content	Number of using graphs
8	14th Apr	Review for mid-term examination	0
9	21st Apr	Mid-term examination	-
10	28th Apr	8-6 Estimation of difference of two population mean $\mu_1 - \mu_2$	0
11	5th May	8-6 Estimation of difference of two population mean $\mu_1 - \mu_2$ 8-7 Estimation of difference of two population mean $p_1 - p_2$	0
12	12th May	9-1 Basic ideas: idea of hypothesis testing, types of hypothesis	2
13	19th May	9-1 Basic ideas: type I error and type II error	6
14	26th May	9-1 Basic ideas: how to decide α 9-2 Testing population mean μ	5
15	2nd Jun	9-3 Testing population proportion p 9-5 Testing difference of two population mean $\mu_1 - \mu_2$	2
16	9th Jun	9-6 Testing difference of two population proportion $p_1 - p_2$	0
17	16th Jun	Review for final examination	0
18	23rd Jun	Final examination	-

Table 6.1. Number of graphs used each week.

**If one graph was repeatedly used by the teacher to explain different ideas, the number of graphs used was counted as multiple times.*

I noticed some of the teacher's teaching problems during the twelfth and fifteenth weeks.

Since the teacher used a similar teaching strategy and process when he did not use graphs, I felt it necessary to investigate the teacher's teaching problems that occurred during the four weeks. According to the proposed research questions (see Section 1.2), the focus of the observation was on the teacher's teaching strategy and the application and influence of his graphs. I investigated the teacher's teachings in hypothesis testing and designed exercise questions to test whether the students had such or other problems. I also explored the causes of the teacher's problems in the second teacher interview. From Section 6.3.3.1 to Section 6.3.3.7, I will illustrate the teacher's observed problems which might influence the students. I will also explain how I designed questions to explore these issues.

6.3.3.1 Critical value not linked to the type of testing

When teaching hypotheses testing in the twelfth week, the teacher drew a graph of a normal curve with two tails (see Figure 6.2) and told the students that the critical value was relevant to the significant level. However, he did not mention that the critical value was also relevant to the type of testing, such as how two-tailed testing and one-tailed testing led to different αs, and how left-tailed testing led to a negative critical value. That he did not consider the type of testing will be further discussed in Section 6.3.3.4, Section 6.3.3.5 and Section 6.3.3.6.

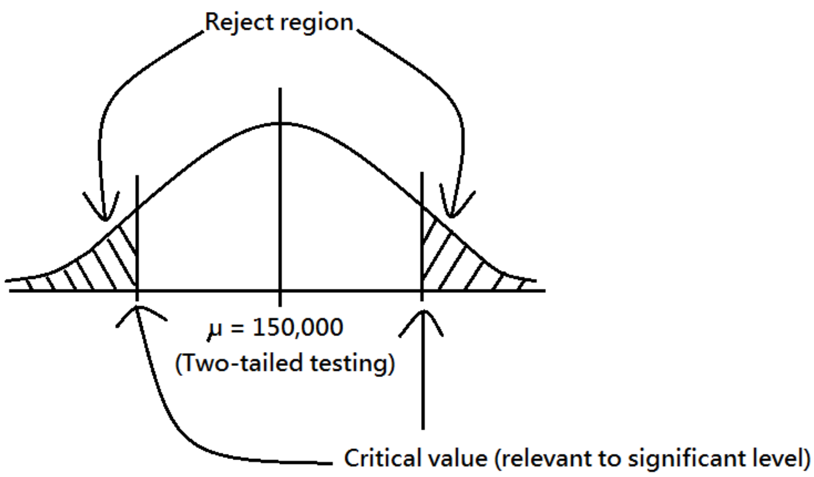


Figure 6.2. Graphical representation of reject regions and critical values in the

teacher's example.

6.3.3.2 Unmatched hypothesis and graph

When introducing the idea of type I error α in the thirteenth week, the teacher drew a curve of H_0 and wrote down the two-tailed hypotheses $\begin{matrix} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{matrix}$. However, he drew only the left-tail rather than the two-tails below the curve (see Figure 6.3.). That is, the teacher's graph did not match his hypotheses. I explored the reason why the teacher's graph did not match the hypotheses and examined whether the students had such problems in the clinical interviews.

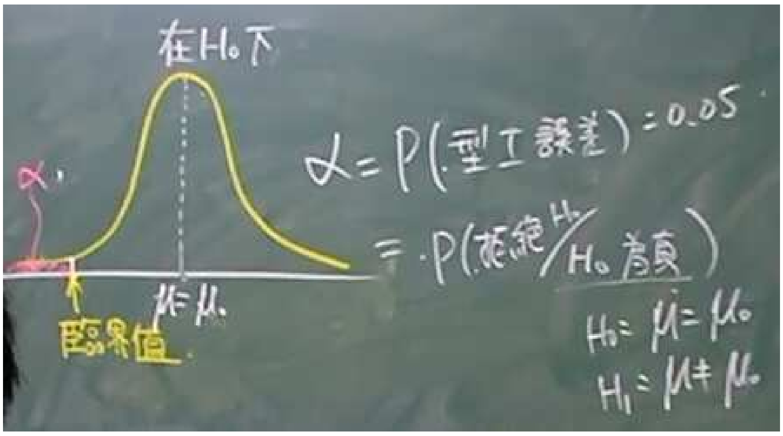


Figure 6.3. The teacher's graph of the left-tailed curve, which did not match his two-tailed hypotheses.

6.3.3.3 Memorised critical value without explanation of its origin

With the graph of Figure 6.3, the teacher explained the idea and use of type I error α:

When α is known, it is possible to find the critical value;

The found critical value is compared to the test statistic which can be obtained by putting numbers in the formula and calculation; and

If the test statistic falls in the region of α (reject region), then H_0 should be rejected.

However, the teacher asked the students to memorise some critical values corresponding to the given α without explaining again how to find them in the tables of distribution according to α. Thus, I examined whether the students understood the meaning of α in the tables of normal distribution and whether they were able to find the critical values on their own. I also noticed any possible problems and explored the causes.

6.3.3.4 Confusion of critical value and confidence interval

In an exemplar question, the teacher pointed at the graph with the left-tail (see Figure 6.4.) and used the memorised "z score -1.96 for $\alpha = 0.05$" as the critical value. In this case, he did not consider that his graph was left-tailed, and he appeared to be misled by the memorised z score of confidence interval. The teacher said:

> Significant level 0.05, confidence degree 0.95, so the z score is 1.96. Remember that? (19th May 2009).

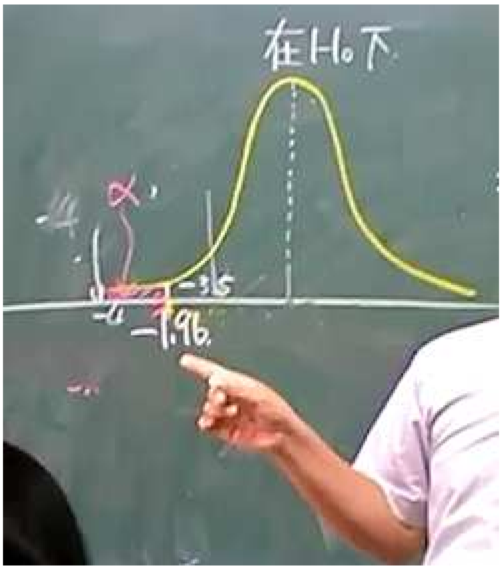


Figure 6.4. The teacher recalled -1.96 according to $\alpha = 0.05$.

As discussed in Section 6.3.3.3, the teacher did not review how to obtain the z scores of confidence interval in the tables of normal distribution and simply asked the students to memorise 1.645 for 90%, 1.96 for 95% and 2.575 for 99% confidence interval. However, he did not consider that the critical value of α = 0.05 of one-tailed testing is not 1.96 and that the critical value of left-tailed testing was not positive. Some students might not be able to distinguish the difference. Thus, I decided to test whether the students could find the correct z scores according to the given probability before/without recalling the memorised z scores.

Some textbook authors might have noticed such confusion in students. For instance, Bluman (2008, front appendix) includes ‘confidence intervals’ and both ‘one tail, α’ and ‘two tails, α’ in the table of t distribution (see Table 6.2.). Such a table helps the readers to directly find the t score according to the given probability and condition without calculation of probability when dealing with most questions of confidence interval and hypothesis testing.

Table F The *t* Distribution

	Confidence intervals	50%	80%	90%	95%	98%	99%
	One tail, α	**0.25**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
d.f.	**Two tails, α**	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
1		1.000	3.078	6.314	12.706	31.821	63.657
2		.816	1.886	2.920	4.303	6.965	9.925
3		.765	1.638	2.353	3.182	4.541	5.841
4		.741	1.533	2.132	2.776	3.747	4.604
5		.727	1.476	2.015	2.571	3.365	4.032
6		.718	1.440	1.943	2.447	3.143	3.707
7		.711	1.415	1.895	2.365	2.998	3.499
8		.706	1.397	1.860	2.306	2.896	3.355
9		.703	1.383	1.833	2.262	2.821	3.250
10		.700	1.372	1.812	2.228	2.764	3.169
11		.697	1.363	1.796	2.201	2.718	3.106
12		.695	1.356	1.782	2.179	2.681	3.055
13		.694	1.350	1.771	2.160	2.650	3.012
14		.692	1.345	1.761	2.145	2.624	2.977
15		.691	1.341	1.753	2.131	2.602	2.947
16		.690	1.337	1.746	2.120	2.583	2.921
17		.689	1.333	1.740	2.110	2.567	2.898
18		.688	1.330	1.734	2.101	2.552	2.878
19		.688	1.328	1.729	2.093	2.539	2.861
20		.687	1.325	1.725	2.086	2.528	2.845
21		.686	1.323	1.721	2.080	2.518	2.831
22		.686	1.321	1.717	2.074	2.508	2.819
23		.685	1.319	1.714	2.069	2.500	2.807
24		.685	1.318	1.711	2.064	2.492	2.797
25		.684	1.316	1.708	2.060	2.485	2.787
26		.684	1.315	1.706	2.056	2.479	2.779
27		.684	1.314	1.703	2.052	2.473	2.771
28		.683	1.313	1.701	2.048	2.467	2.763
(z) ∞		.674	1.282[a]	1.645[b]	1.960	2.326[c]	2.576[d]

[a]This value has been rounded to 1.28 in the textbook.
[b]This value has been rounded to 1.65 in the textbook.
[c]This value has been rounded to 2.33 in the textbook.
[d]This value has been rounded to 2.58 in the textbook.

Source: Adapted from W. H. Beyer, *Handbook of Tables for Probability and Statistics*, 2nd ed., CRC Press, Boca Raton, Florida, 1986. Reprinted with permission.

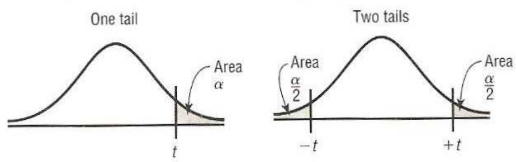


Table 6.2. The table of t distribution in Bluman's (2008) textbook.

Bluman (2008) adds both a right-tailed graph and a two-tailed graph below the table (see Figure 6.5.). This kind of table and graph pairing helps prevent readers from getting confused about the given probability (such as 90%, 95% or 99%) and α (such as 0.05). However, many textbooks, including the textbook used in the observed course, do not use such tables and graphs, so it is necessary to take whether the students consider the α issue into account.

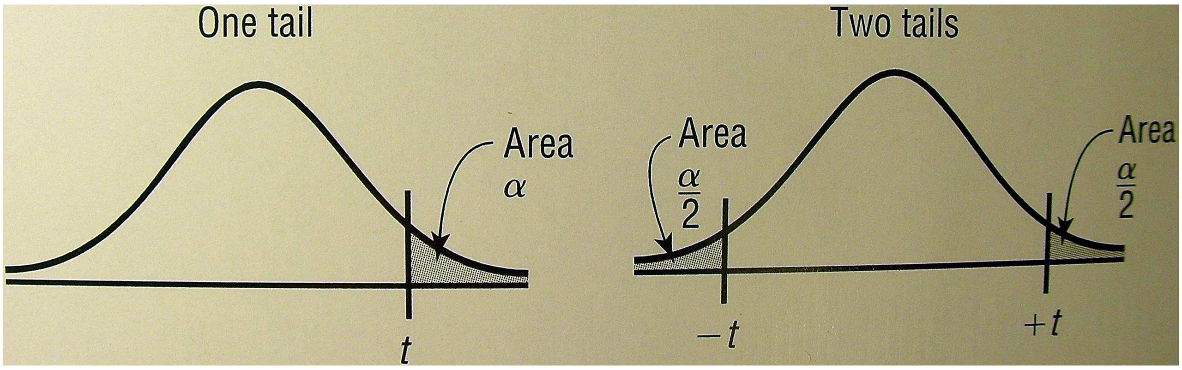
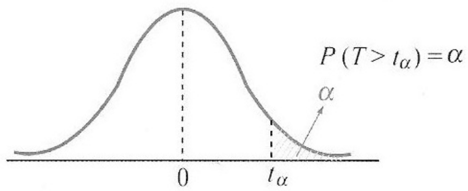


Figure 6.5. The graphs of t curve with shadow in right tail and with shadows in two tails were attached below the table of t distribution.

6.3.3.5 When the given probability is not the α to be found in the table

Moreover, the probabilities given in the questions of confidence interval or hypothesis testing were not necessarily the α to be found in their table. For example, the students in the study had to find the correct α according to the given probability and condition before looking for α in the table of t distribution in their textbook (see Table 6.3. and Table 6.4.). I had tested whether and how the participants found the correct α before using the table of t distribution in Q3.1 and Q3.2. Here I tested whether the participants obtained the correct α of one-tailed testing and α of two-tailed testing according to the given α.

表 A–5　t 分配



v	α						
	.25	.2	.1	.05	.025	.01	.00833
1	1.000	1.376	3.078	6.314	12.706	31.821	38.190
2	.816	1.061	1.886	2.920	4.303	6.965	7.469
3	.765	.978	1.638	2.353	3.182	4.541	4.857
4	.741	.941	1.533	2.132	2.776	3.747	3.961
5	.727	.920	1.476	2.015	2.571	3.365	3.534

Table 6.3. The upper-left part of the table of t distribution in the participants' textbook.

表 A-5（續） t 分配

v	.00625	.005	.0025	.001	.0005	.00025	.0001	.00005
1	50.923	63.657	127.21	318.309	363.619	1,273.239	3,183.099	6,366.198
2	8.860	9.925	14.089	22.327	31.598	44.705	70.700	99.992
3	5.392	5.841	7.453	10.214	12.924	16.326	22.204	28.000
4	4.315	4.604	5.598	7.173	8.610	10.306	13.034	15.544
5	3.810	4.032	4.773	5.893	6.869	7.976	9.678	11.178

Table 6.4. The upper-right part of the table of t distribution in the participants' textbook.

6.3.3.6 Type of testing

When introducing the p-value method in the fourteenth week, the teacher explained that the p-value method is used to make a decision in terms of whether the probability is smaller than the significant level. Again, he did not consider whether the test was two-tailed or left/right-tailed. Therefore, I examined whether the students took the type of test into account and understood how to find the probability values.

6.3.3.7 In-complementary hypotheses and the memorised z score

In an exemplar question of the textbook, the teacher was observed making the hypotheses $\begin{array}{l} H_0 : \mu = 100 \\ H_1 : \mu > 100 \end{array}$ in terms of $\overline{X} = 106$. His hypotheses were the same as those in the textbook, which might confuse the users as to whether the testing was two-tailed or right-tailed. He made the critical region CR = {Z > $z_{0.05}$} and used the memorized z score 1.96 for $\alpha = 0.05$ again (see Figure 6.6).

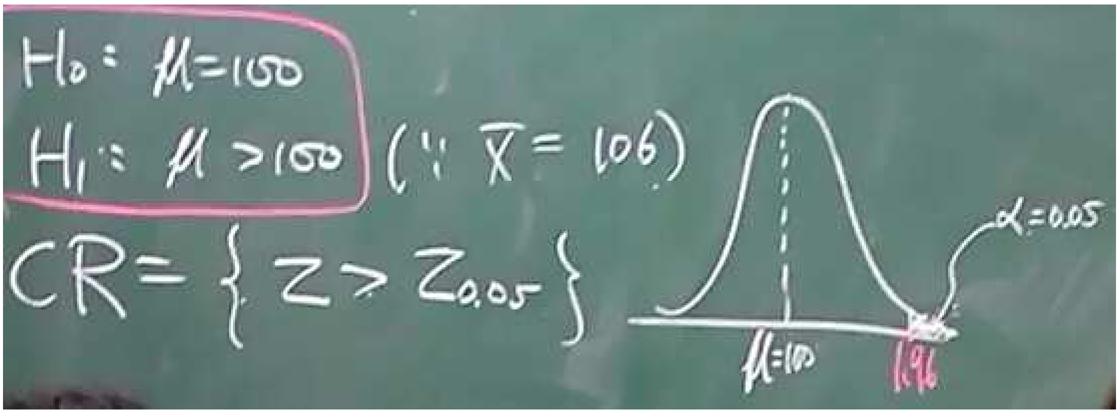


Figure 6.6. The teacher's graph and wrong z score in Q9-1.13.

The teacher repeated similar problems (since the thirteenth week) until some students noticed that the teacher's z score was different from that in the textbook one week later. The students suggested that the z score should be 1.645 (perhaps according to the solution in the textbook), and after considering this for awhile, the teacher noticed his problem. The teacher pointed at the shadowed right-tail and explained that 0.05 is the right-hand side area of 1.645 and 0.025 is the right-hand side area of 1.96 (see Figure 6.7.).

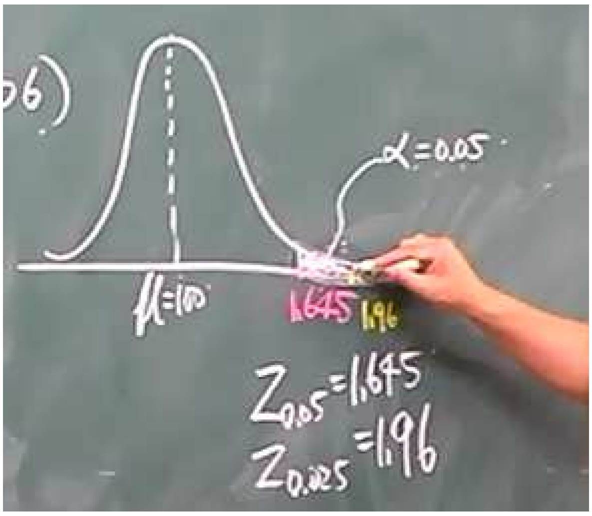


Figure 6.7. The teacher explained $z_{0.05} = 1.645$ *and* $z_{0.025} = 1.96$ *in the graph.*

The teacher's repeated problem implies the importance of exploring whether the participants were able to make complementary hypotheses (at least distinguishing the type of testing) and find correct critical values or simply use the recalled (positive) z scores of two-tailed confidence interval regardless of the type of testing (two-tailed, left-tailed or right-tailed), and whether their hypotheses influenced their choice of critical value. I also examined whether the participants were able to carry out the process of hypothesis testing and make decision. Therefore, I designed two hypothesis testing questions with similar conditions in the second student interview to explore whether the participants distinguished the questions as left-tailed, right-tailed or two-tailed and make correct hypotheses according to the conditions that ‘the content is 500cc’ and ‘the content is over 500cc’. These two questions will allow me to see how the participants find the critical area(s), what problems they have, and whether/how they used graphs in the solving processes. This will be discussed in Chapter 10.

6.4 Teacher interview II (week 16)

The second teacher interview was conducted one week after the second student interviews. At the beginning of the interview, I let the teacher review the results of the students’ exercises including the correct/wrong/no answer ratio, the commonly wrong answers and my initial analysis of their solutions, and I invited him to discuss the reason(s) for the participants’ difficulties. I also let the teacher explain his problems with teaching hypothesis testing, such as the recalled z score.

The teacher noticed the students’ difficulty with questions involving negative value and symmetric ideas and their confusion with bigger than and smaller than. Thus, he suggested that some students might lack the concept of ‘direction’ and were not concerned with the relation signs, and that visualisation would help in dealing with such problems. Unfortunately, after the discussion of the results in the student interview, the teacher had to leave, and he was not able to attend another interview due to his tight schedule. However, I felt that the teacher seemed to become worried about his lack of experience in teaching statistics and the teaching problems identified by a colleague. This might be another reason that he decided not to take part in further interviews.

6.5 Student interview III (week 18)

Five participants (S1-S5) who had not prepared well for the second interview were interviewed a third time after the final examination. The main issue of the third interview was to test how they distinguished and solved hypothesis testing questions, and I also tested the two similar hypothesis testing questions used in the second interview (i.e. Q6 and Q7) again at the end of the third interview as Q9 and Q10. Before these two questions, I asked the participants to review the whole course, explain their general strategies of question-solving, express their attitude towards the course and the teacher, and offer some suggestions to the teacher. I also invited them to talk about their thoughts about the knowledge and use of statistics.

6.6 Conclusion

After the second semester, the audio- and video-recordings of the second and third student interviews were also transcribed, listed, coded and categorised in Appendix 6.20. – Appendix 6.22.

Appendix 6.22 extended the scope of Appendix 6.10. and provided more details of the key ideas of learning hypothesis testing, difficulties and causes, students' strategies and students' interpretation of important ideas. These findings were the basis of the study and led me to construct the chapters of data analysis in Chapter 7 – Chapter 11.

CHAPTER 7

DATA ANALYSIS I

APPLICATION OF THE TABLES OF STANDARD NORMAL DISTRIBUTION

7.1 Introduction

The content taught during the data collection period was the part of statistical inference. In the first week of data collection, the teacher reviewed the content which he had taught in the first seven weeks of the second semester as the preparation for the mid-term examination. This review included sampling distribution and estimation. After the mid-term, the teacher taught estimation of difference of two means $\mu_1 - \mu_2$ and estimation of difference of two proportions $p_1 - p_2$, as well as hypothesis testing of mean μ, of proportion p, of difference of two means $\mu_1 - \mu_2$ and of difference of two proportions $p_1 - p_2$. However, he excluded the estimation and hypothesis testing of variance because of its difficulty. He also excluded χ^2 distribution and F distribution. According to the content taught in the class, I designed exercise questions to examine how students learned in this semester. The use of tables of normal distribution and of t distribution was taught in the first semester by another teacher, but these concepts were tested because they were still frequently used in this semester. Therefore, the exercise questions consisted of four categories, which are

I. Application of the tables of standard normal distribution;

II. Application of the table of t distribution;

III. Definition of confidence interval andhow to find the z score of confidence interval; and

IV. Hypothesis testing, including finding the critical value.

I will analyse the participants' solutions to the exercise questions in these four categories in Chapters 7, 8, 9 and 10 respectively. I will also discuss whether and how the participants used or did not use visualisation and the influence on their solutions in Chapter 11. In this chapter, I will discuss how the participants used the tables of normal distribution to solve Q1.1-Q1.6 in the first interview (see Appendix 7.1.) and Q1.a-Q1.d in the second interview (see Appendix 7.2.):

Please look for the value in the table of normal distribution:

Q1.1 $P(Z < -0.03)$;

Q1.2 $P(Z < 1.22)$;

Q1.3 $P(Z > -3.1)$;

Q1.4 $P(Z > 2.05)$;

Q1.5 $P(-1.4 < Z < 2.32)$; and

Q1.6 $P(-2.41 < Z < -0.03)$

and

Q1.a $P(Z > a) = 0.67$, find a;

Q1.b $P(Z > 0.67) = b$; find b;

Q1.c $P(Z < c) = 0.67$, find c; and

Q1.d $P(Z < 0.67) = d$, find d

Although there have been many studies on students' learning and difficulties with statistical reasoning, the issue of 'how students realise and use the tables of distribution' seems to have been overlooked by previous research. In order to investigate this question, I asked the participants at the beginning of the first interview to find the probability according to the given relation sign and z score in Q1.1-Q1.6, and to find the t score according to the given probability and degree of freedom in Q3.1-Q3.2 (see Chapter 8). The tables of normal distribution and the table of t distribution were tested because they are "the two most useful tables in statistics" (Wonnacott and Wonnacott, 1990, p.262). According to the findings from the first interview, I asked the participants to find the value in the tables of normal distribution in terms of the information given by questions in different formats in the second interview.

In order to give the readers a clearer view of the participants' difficulties, I will first introduce seven symptoms of difficulties observed in their solutions by comparing and contrasting each participant's solution in each question to the correct solution and his/her solutions in other questions in order to make sure each problem encountered by the participants was considered. I will then suggest several factors which may have caused the participants' difficulties and explain how these factors influenced the participants' solutions. The participants had several kinds of solutions, and I will summarise the solutions with similar problems and give a few representative examples.

Before examining the participants' difficulties, I note that while most participants were interviewed in groups, each of them solved most of the exercise questions on his/her own. However, some participants repeated others' explanations when discussing the solving strategies, and some participants asked others to explain particular ideas when solving some questions (such as in Chapter 9). For such situations, I will include the participants' interaction and the influence in the analysis.

7.2 Symptoms of difficulties

When the participants solved the exercise questions in the interviews, they had many problems, some of which were caused simply by being careless. Such carelessness in question-solving was found as early as in 1959, when Maxwell defined a mistake as being caused by “a momentary aberration, a slip in writing or the misreading of earlier work” (1959, p.9). Different researchers may use different terminologies to express the same idea, and they may also use the same terminology to represent different ideas. For example, Elbrink (2008) refers to mistakes caused by carelessness and lack of attention as “calculation errors” (2008, p.2). She also defines “procedural errors” and “symbolic errors” (2008, p.2) in students’ question-solving. Legutko (2008) distinguishes between a mistake and an error as follows:

> A *mistake* is a result of the lack of concentration control or weak memory. We make a mistake when we incorrectly apply a formula or theorem, which we know (or should know) from theory acquired earlier. An *error* reveals inadequacy of knowledge and is closely connected with imagination and creativity in a new situation, and is caused by an insufficient mastery of basic facts, concepts and skills (Legutko, 2008, p.143).

Legutko (2008) adds that an error “takes place when a person chooses the false as the truth” (2008, p.142), which results “from erroneous conceptions” (2008, p.144). To create a sensible error typology is impossible and unnecessary, but it is useful to analyse the causes of errors and to apply the analysis. In order to avoid misunderstanding, I will use the term ‘careless mistakes’ to represent mistakes caused by being careless in this study. Many careless mistakes were found ‘mixed’ with mistakes caused by other factors in the participants’ solutions, and this situation makes it difficult to analyse their answers. Thus, I will point out the careless mistakes in several subcategories here. In the order of possible appearance when solving questions of using the tables of distribution to find a value, I distinguish the problems found in students' responses into seven symptoms, namely mis-reading, mis-transcribing, mis-fitting, mis-transferring, mis-targeting, mis-interpreting and mis-calculation.

Mis-reading the question and the following steps in one's own hand-writing means that a student might not read the given information in the question or the hand-writing correctly. For example, he/she may not notice the existence of the minus sign and thus treat a given negative number as positive, or regard 'bigger/smaller than' as 'smaller/bigger than'. This problem will change the meaning of the question and lead the student to solve a different question. Such a problem occurred in Q1.3, in which S1 and S5 did not see the minus sign in the question and S2 ignored the minus sign in his own hand-writing. Misreading the minus sign and relation sign also happened to S4. The cause of such a problem may not be carelessness, as I will explain in Section 7.3.1.1.

Mis-transcribing means that participants might write different numbers or symbols from those which were in the text of the question or in the previous solving steps, or that they may overlook some of them. For instance, S3 transcribed -0.03 as -0.13 in Q1.6 and lost the P when writing $P(Z > 0.13)$. S7 wrote $(P > 0.03)$ instead of $P(Z > 0.03)$. S10 also lost the P in two questions. Some kinds of mis-transcription did not influence the answers, while other kinds did. These situations result from carelessness.

Mis-fitting means that some participants used the table of distribution when the format of the equation did not fit the format of the table of distribution without or even after transferring steps. For instance, in Q1.2, which involved 'Z being smaller than a positive z score', S2 found the given z score outside Table B, which was about 'Z being bigger than a positive value'. He also found the absolute value of the given z score outside Table B in Q1.3, which involved 'Z being bigger than a negative z score'. The causes of such symptoms deserve further investigation and will be discussed in 7.3.1.1.

Mis-transferring the equation means that the equation was transferred to another written form incorrectly. For instance, S8 transferred P(Z < 1.22) to P(Z > 1.22) in Q1.2 and transferred P(Z > -3.1) to P(Z > 3.1) in Q1.3. Although S2 had the same answers as S8 in Q1.2 and Q1.3, he did not write this transferring step. Therefore, I will not include his solutions in this category. S9 also mis-transferred a 'smaller than' symbol to a 'bigger than' symbol. S7 did not have such problems in basic questions, but she mis-transferred negative z scores to positive ones in Q1.5 and Q1.6. These difficulties may not be merely the result of being careless, and this will also be discussed in Section 7.3.1.1.

Mis-targeting the value in the table of distribution means that the students might use the wrong table or find wrong value(s) when looking for a particular target. For example, S1 used the table of t distribution when he was solving a question of normal distribution. S5 frequently took the left-hand side 'neighbour' of the correct answer as his answer (see Table 7.1.). Mis-targeting seemed to be his biggest problem, since this appeared in four of the six questions. This problem only happened once in another participant's answers.

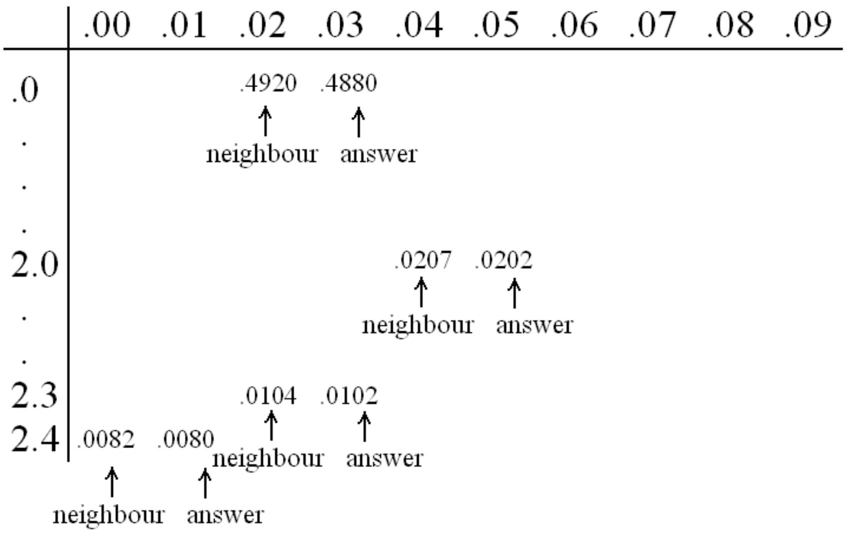


Table 7.1. S5's answers and correct answers.

Mis-interpreting the table means that students might give a different meaning to the found value in the tables of distribution. For instance, S1 mis-interpreted the z score .14 with .1 in the left-hand side and .04 on top of Table B as 1.4, which is in fact located at 1.4 on the left-hand side and .00 on top of the table; S2 found the probability value .0202 inside Table B and treated it as 0.202; S4 found the z score 0.44 (.4 in the left-hand side and .04 on top of the frame of Table B) and regarded it as 0.04 or 0.404; and S10 regarded the value combined of -2.0 on the left-hand side and .04 on the top of Table B as the wanted z score –2.41. Another problem was that students did not distinguish what was outside the table and what was inside the table, and they used the table in a reverse way. For example, S4 and S5 found the given probability 0.67 outside Table B in Q1.a, while it should be included inside the table.

Mis-calculation was also found in many of the participants' solutions. The definition of mis-calculation in my study is limited to the arithmetical operations. Most of the participants' mis-calculations happened to subtraction of decimals and 1. S1, S3 and S9 mis-calculated when they subtracted decimal(s) from 1, as in the following examples:

1 – 0.0102 = 0.898 (by S1);

1 – 0.0808 – 0.0104 = 0.0088 (by S3); and

1 – 0.0010 = 0.9989 (by S9).

Another kind of mis-calculation was "loss of invisible parentheses" (Schechter, 2009). For example, S5 solved Q1.7 in this way:

$$
\begin{aligned}
&P(Z > -2.41) - P(Z > -0.03) \\
&= 1 - P(Z > 2.41) - 1 - P(Z > 0.03) \\
&= 1 - 0.0082 - 1 - 0.4880
\end{aligned}
$$

whereas the correct answer should be '$1 - 0.0082 - 1 +\ 0.4880$'. However, such problems did not necessarily influence the result if it was a "cancelling error" (Crowley and Tall, 2006, pp.63-64), such as a part of S9's solution in Q1.7:

$$
\begin{aligned}
&P(x > -2.41) - P(x > -0.03) \\
&= 1 - P(x > 2.41) - 1 - P(x > 0.03) \\
&= (1 - 0.008) - (1 - 0.4880)
\end{aligned}
$$

Therefore, the missed parentheses in the second row were put back in the third row, and the result was not influenced.

Numerical and logical mistakes were found in Q1.1-Q1.6 and Q3.1-Q3.2 in the first interviews, as well as in Q1.a-Q1.d in the second interviews. The frequencies of the seven symptoms of problematic solution which occurred in Q1.1-Q1.6 and Q1.a-Q1.d are summarised in the upper part of Table 7.2 and Table 7.3 respectively (for the details, see Appendix 7.3.).

Symptoms & Causes (Stage)			Question	Infor-mation	Inter-pretation	Action
Symptoms	Mis-reading		-	-	2	-
	Mis-transcribing		-	-	-	6
	Mis-fitting		-	-	5	0
	Mis-transferring		-	-	-	8
	Mis-targeting		-	-	-	5
	Mis-interpreting		-	-	-	4
	Mis-calculation		-	-	-	5
Causes	Reading the question in symbolic form	Relation sign	-	-	4	3
		Minus sign	-	-	8	4
		When Z is between two values	-	-	-	11
		Question format	-	-	4 + 8	7
	Using the table(s)		-	-	-	9
	When the answer was not in the table		-	-	9	9

Table 7.2. Frequencies of symptoms and causes in participants' solutions in Q1.1-Q1.6.

** Numbers in the table mean how many times such a situation happened; repeated symptoms in the same question were counted multiple times; '-' means inappropriate.*

***The counting of the numbers is explained in Appendix 7.3.*

The causes of the symptoms are listed in the lower part of Table 7.2 and Table 7.3 (for the details, see Appendix 7.3.). The causes are obtained by analysing the participants' solutions and categorising them into three general factors, namely reading the question in symbolic form, using the table, and when the answer is not in the table. The details of the causes will be discussed in Section 7.3.

<table>
<tr><th colspan="3">Stage
Symptoms & Causes</th><th>Question</th><th>Infor
-mation</th><th>Inter
-pretation</th><th>Action</th></tr>
<tr><td rowspan="7">Symptoms</td><td colspan="2">Mis-reading</td><td>-</td><td>-</td><td>0</td><td>-</td></tr>
<tr><td colspan="2">Mis-transcribing</td><td>-</td><td>-</td><td>-</td><td>0</td></tr>
<tr><td colspan="2">Mis-fitting</td><td>-</td><td>-</td><td>4</td><td>0</td></tr>
<tr><td colspan="2">Mis-transferring</td><td>-</td><td>-</td><td>-</td><td>4</td></tr>
<tr><td colspan="2">Mis-targeting</td><td>-</td><td>-</td><td>-</td><td>4</td></tr>
<tr><td colspan="2">Mis-interpreting</td><td>-</td><td>-</td><td>-</td><td>4</td></tr>
<tr><td colspan="2">Mis-calculation</td><td>-</td><td>-</td><td>-</td><td>0</td></tr>
<tr><td rowspan="6">Causes</td><td rowspan="4">Reading the question in symbolic form</td><td>Relation sign</td><td>-</td><td>-</td><td>6</td><td>3</td></tr>
<tr><td>Minus sign</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>When Z is between two values</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>Question format</td><td>-</td><td>-</td><td>17</td><td>5</td></tr>
<tr><td colspan="2">Using the table(s)</td><td>-</td><td>-</td><td>-</td><td>8</td></tr>
<tr><td colspan="2">When the answer was not in the table (Negative answer)</td><td>-</td><td>-</td><td>3</td><td>1</td></tr>
</table>

Table 7.3. Frequencies of symptoms and causes in participants' solutions in Q1.a-Q1.d.

* *Numbers in the table mean how many times such a situation happened; repeated symptoms in the same question were counted multiple times;'-' means inappropriate.*

***The counting of the numbers is explained in Appendix 7.3.*

In Table 7.2 and Table 7.3, we can see that the participants' responses were often problematic. However, careless mistakes will not be addressed in my analysis, and the focus of analysis will be on how the participants solved the exercise questions and whether they followed the correct solving steps. That is, a participant who made careless mistakes in finding an answer will not be investigated unless he/she went through incorrect solving steps. In this way, we can focus on the participants' problems with solving steps. Below, I introduce the three general factors which caused the participants' difficulties and discuss how each factor influenced participants' solutions.

7.3 Three factors of difficulties

7.3.1 Reading the question in symbolic form

In the analysis of participants' solutions in the interviews, a difficulty was found to be caused by the participants' interpretation of the questions in symbolic form. Unlike other subjects such as physics, geology and botany, using symbolic language is unavoidable for mathematical thinking since

> there is no other ways of gaining access to the **mathematical objects** but to produce **some semiotic representations**… [and] the understanding of mathematics requires not confusing the mathematical objects with the used representations (Duval, 1999, p.4).

Hilbert *et al.* (1997) explain the difficulty of understanding symbolism, as "meaning is not inherent in symbols" (1997, p.55). They continue that symbols cannot be used effectively without meaning. Corter and Zahner (2007) suggest that solving problems always involves four steps:

> initial problem understanding (text comprehension);
> formulating the mathematical problem;
> finding a solution or schema; and
> computing the answer (2007, pp.25-26).

Since the questions involving the use of the tables of normal distribution (i.e. Q1.1-Q1.6 and Q1.a-Q1.d) were presented in symbolic form (i.e. the second stage), the participants only needed to find the solution or schema by doing some necessary computations. Riley *et al.* (1983) find that failures to solve problems are often caused by a lack of appropriate schemas, and problem solvers frequently correct arithmetic process on incorrect representation of the problems (1983, cited in Corter and Zahner, 2007). Therefore, the schemas to solve the questions of using the tables of normal distribution will be the main interest in this study.

Here I will first explain how the two elements of questions in symbolic form, relation sign and negative sign, influenced the participants' question solving in Q1.1-Q1.6. Although the two elements are different, I will discuss them together because when the participants made a mistake with one element in a question, they always made a mistake with the other element in the same or other questions. This will be followed by analysis of how the question of z between two values in Q1.5 and Q1.6 and the different formats of questions in Q1.a-Q1.d caused the participants' difficulties.

7.3.1.1 Relation sign and minus sign

As explained in Chapter 5, when the question format matched the formats of the chosen table of normal distribution, the users could simply find the answers in the table. Only Q1.4, which matched the format of Table B, could be solved without any transferring. There was no Table B user who performed unnecessary transferring or made mistakes in this question. When the question format does not match the format of the chosen table, transferring steps should be applied to transfer the question to a suitable format before using the tables. However, some transferring procedures undertaken by the participants were meaningless, and the equilibrium of the equation was not maintained after transferring.

When the participants solved the basic questions Q1.1-Q1.4, not all of them correctly interpreted or even noticed the information given in the question. In the first interview, S4 used Table A, and the other nine participants used Table B. No matter which table was used, there were two kinds of symbol that the participants made mistakes with: the relation sign and the minus sign (see Table 7.4).

Question	Q1.1	Q1.2	Q1.3	Q1.4
	P(Z < -0.03)	P(Z < 1.22)	P(Z > -3.1)	P(Z > 2.05)
Number of Table A users who applied the correct approach in Q1.1-Q1.4	0	0	0	0
Number of Table A users who applied the incorrect approach in Q1.1-Q1.4	1	N/A (no answer)	1	1
Number of Table B users who applied the correct approach in Q1.1-Q1.4	5	6	8	9
Number of Table B users who applied the incorrect approach in Q1.1-Q1.4	4	3	1	0
Causes of incorrect approach	Relation sign: S1, S2, S8, S9; Minus sign: S4, S1, S2, S8, S9	Relation sign: S2, S8, S9	Relation sign: S4 Minus sign: S2, S8	Relation sign: S4

Table 7.4. Number of Table B users who applied the correct approach and incorrect approach in Q1.1-Q1.4 and the causes of incorrect approach.

**In Q1.1, S1 might have read P(Z < -0.03) as P(Z < 0.03), which involved a minus sign, or transferred P(Z < -0.03) to 1 - P(Z > 0.03) because the original relation sign did not match the format of Table B. Which step was taken by S1 cannot be confirmed, so both possible causes are listed here with explanation.*

In this section, I will use three less successful students, S2, S8 and S4, as examples. Their solutions in Q1.1-Q1.6 were summarised with the correct answers in Table 7.5. An analysis of their solutions is added below to explain how they solved the questions.

	Q1.1	Q1.2	Q1.3	Q1.4	Q1.5	Q1.6
	P (Z < -0.03)	P (Z < 1.22)	P (Z > -3.1)	P (Z > 2.05)	P (-1.4 < Z < 2.32)	P (-2.41 < Z < -0.03)
Correct Answer	**0.488**	**0.8888**	**0.999**	**0.0202**	**0.9898 - 0.0808 = 0.909 or 1 – 0.0808 – 0.0102**	**0.488 – 0.008 = 0.48 or (1 - 0.008) - (1 - 0.488)**
S2's answer	-0.03; 0.4880	< 1.22; 0.1112	> - 3.1; 0.010	> 2.05; 0.202	-1.4 < Z < 2.32; 0.0808 < Z < 0.0102	-2.41 < Z < -0.03; 0.0080 < Z < 0.4880
Analysis of S2's answer	P(Z > 0.03) = 0.488	P(Z > 1.22) = 0.1112	P(Z > 3.1) = 0.001	P(Z > 2.05) = 0.0202	P(Z > 1.4) = 0.0808 P(Z > 2.32) = 0.0102	P(Z > 2.41) = 0.0080 P(Z > 0.03) = 0.4880
S8's answer	P (Z < -0.03) = P (Z > 0.03) = 0.4880	P (Z < 1.22) = P (Z > 1.22) = 0.1112	P (Z > -3.1) = P (Z > 3.1) = 0.9990	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) =	P (-2.41 < Z < -0.03) =
Analysis of S8's answer	P (Z > 0.03) = 0.4880	P (Z > 1.22) = 0.1112	P (Z > 3.1) = 0.9990	P (Z > 2.05) = 0.0202		
S4's answer	P(Z < -0.03) = 0.5120	P(Z < 1.22)	P(Z > -3.1) = 0.0010	P(Z > 2.05) = 0.9798	P(-1.4 < Z < 2.32)	P(-2.41 < Z < -0.03)
Analysis of S4's answer	P(Z < 0.03) = 0.5120		P(Z < -3.1) = 0.001	P(Z < 2.05) = 0.9798		

Table 7.5. Comparison and analysis of S2's, S8's and S4's solutions of Q.1-Q1.6 to the correct answers.

In the analysis of the participants' answers to the exercise questions, whether a participant was able to use the tables of distribution could not be judged according to whether he/she obtained the correct answer. Batanero *et al.* (1996) remark that

> it is possible to get a correct judgment with an incorrect strategy… the ability to provide both a correct judgment and a correct strategy are needed for an adequate understanding of this concept (p.158).

For instance, although S2 correctly transcribed the minus sign and the smaller than sign in the sub-questions of Q1 (except for Q1.1) on his answer sheet, he did not consider the influence of the 'smaller than sign' in Q1.2 and the 'negative sign' in Q1.3, and he found wrong answers on both questions. He did not consider whether the relation sign and positive/negative z score in each question matched the format of Table B (i.e. Z is bigger than a positive z score), and he treated all relation signs as 'bigger than' and all z scores as 'positive'. He had the correct answers for Q1.1 and Q1.4, but his answer for Q1.1 seemed to be a coincidental result of ignoring both the bigger than and negative signs, rather than a correct solution. Such "mathematical luck" (Aberdein, 2007, p.13) comes from a "'compensation of errors', whereby errors made at different stages systematically cancel each other out" (Aberdein, 2007, p.14).

S2 also failed to consider the minus signs in Q1.5 and Q1.6, and he stopped solving after he found the probabilities corresponding to the mistakenly interpreted positive z scores. He did not present any transferring step in solving the questions, but all of the probabilities in his answer sheet can be found inside Table B if we assume that the z score is positive (i.e. without minus sign) and that the relation sign is 'bigger than'.

Interpreting symbolic equations also seemed to be problematic, and the signs might not be meaningful to students. For instance, Kieran (1981) finds that some students do not interpret the equal sign as an equivalence relation and that they could only build an operational relationship rather than a meaningful link between both sides of the equal sign. Without considering the influence of the relation sign and positive/negative z score (i.e. without/with minus sign), S2 appeared to regard the absolute value of the given z score and the requested probability as a corresponding pair which were collected in Table B. Therefore, when he saw the z score in the question, he directly found the probability value inside Table B corresponding to the z score as the answer.

Unlike S2, who did not consider the relation signs and negative signs of z scores, S8 wrote transferring steps in her answer sheet, but she transferred the questions by replacing 'smaller than' in the questions with 'bigger than' and removing the negative signs. In other words, S8 simply made Q1.1-Q1.4 'become' the basic form for Table B without considering how the changes of sign influenced the equation.

In order to explain and examine S2 and S8's solving steps, it is beneficial to use "visual display" (Harry *et al.*, 2004, p.4) rather than simply narrative description. Therefore, S2's and S8's solutions can be expressed in Figure 7.1.

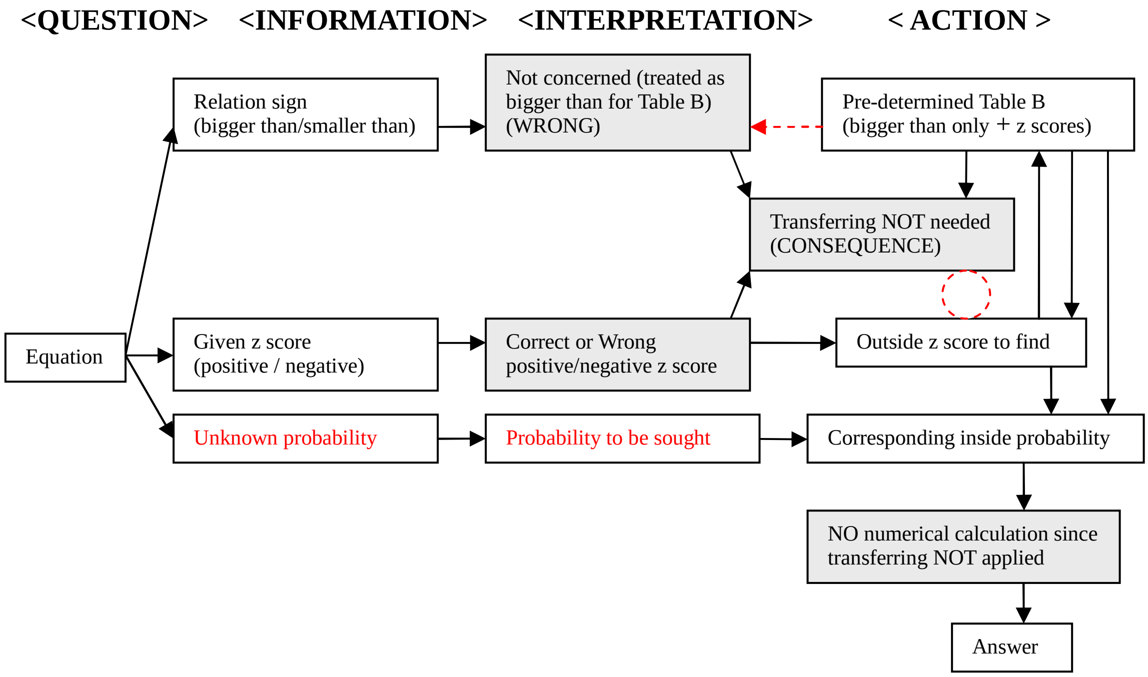


Figure 7.1. S2's and S8's solving steps in Q1.1-Q1.4.

**The dotted line means that S2's and S8's pre-determined Table B caused them not to consider the relation sign in the question; the dotted circle means that the outside z score to find was totally determined by the interpreted z score since transferring was not applied.*

S4 used Table A in the first interview, and she also had problems with the relation sign and minus sign. She did not notice the different relation signs in Q 1.3 and Q1.4 and the existence of the minus sign in Q1. Her approach may be explained by the discussion below:

I: …what is your problem? You did not answer all questions.
S4: I don't understand the formulas.
I: How about the table?
S4: I should be able to use it.
I: In Q1.4, can you show me how to find it?
S4: 0.0202
I: why did you write 0.9798 ten minutes ago?
S4 laughed: Because I use Table A.
I: In Table A, do you just look for 2.05 and ignore smaller than or bigger than?
S4: Yes.
…
I: [in Q1.1,] where is 0.512?
S4 pointed at 0.03: here.
I: Did you notice that it is negative?
S4: No.
(S4, 29th Apr 2009)

Interestingly, S4 did not answer Q1.2, even though the answer to this question could be found directly in Table A and which was supposed to be the simplest question for a Table A user. This finding inspired me to test whether the participants could recognise and distinguish the formats of questions and of the tables of normal distribution in the second interview. I will discuss how question formats influenced the participants' solutions in Section 7.3.1.3.

S4 gave up on Q1.5 and Q1.6. In the only three questions she answered, no transferring was applied since relation signs were not considered and both negative and positive z scores could be found in Table A. S4 regarded the questions as the basic form for Table A: P(Z < some value) = answer, and her solving process of the four basic questions was similar to S2's and S8's approach (see Figure 7.2.). She did not obtain the correct answer to any of the questions.

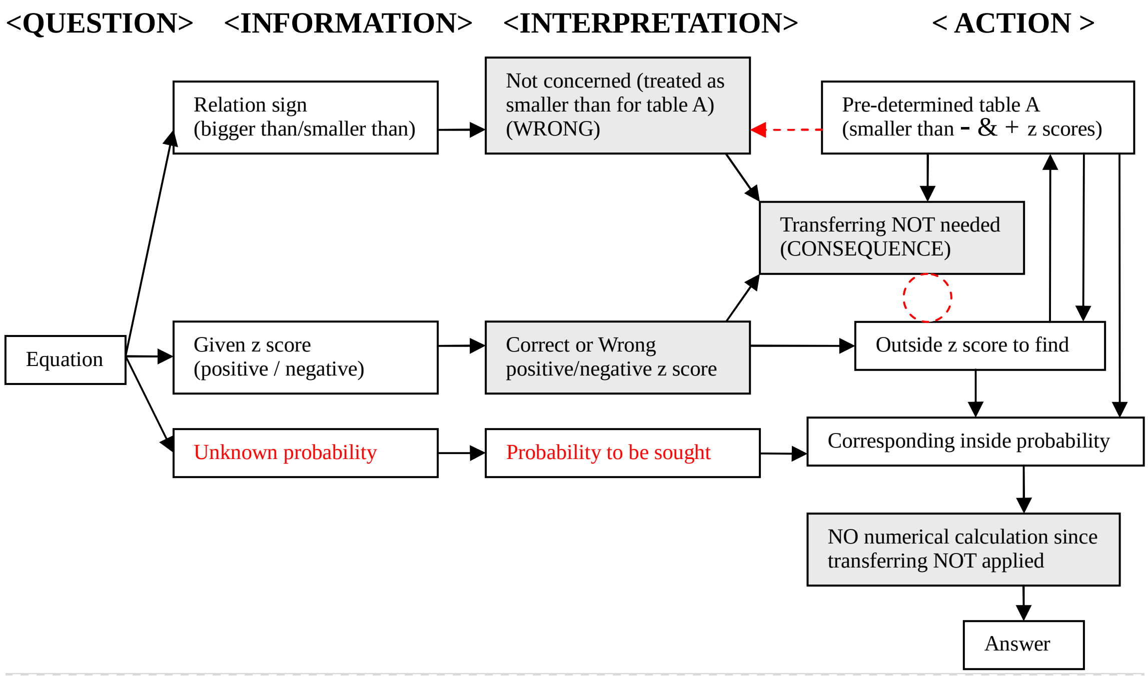


Figure 7.2. S4's solving steps in Q1.1-Q1.4.

**The dotted line means that S4's pre-determined Table A caused her not to consider the relation sign in the question; the dotted circle means that the outside z score to find was decided by the interpreted z score since transferring was not applied*

According to the analysis of the three participants' solutions, we can see that their inattention to the relation symbol and negative sign was influenced by their choice of table. S4 used Table A and treated the two 'bigger than' questions as 'smaller than' questions. Similarly, S2 and S8, both of whom used Table B, treated the 'smaller than' symbols as 'bigger than' and treated the negative value as a positive value. Consequently, they avoided using the transferring principles and found the value in the table(s) as the answer directly.

Although they are not investigated in this study, the fact that S1 and S5 lost the minus sign in Q1.3 when reading the question reflected the gap in students' recognition of negative numbers. This serves as a reminder that educators and researchers need to strengthen students' recognition. In the next section, I will discuss how the participants dealt with the two advanced questions Q1.5 and Q1.6, which could not be solved without transferring steps.

7.3.1.2 When Z is between two values

S4 and S8 did not answer the advanced questions Q1.5 and Q1.6. Five participants transferred each question to two basic formats correctly, while S1 and S2 used a similar method without writing transferring steps and S7 had a confused idea of "betweenness" (Arcavi, 2003, p. 221) in her transferring steps in both questions. S1's, S2's and S7's answers are compared with the correct transferring steps in Table 7.6.

	Q1.5	Q1.6
Question	P(-1.4 < Z < 2.32)	P(-2.41 < Z < -0.03)
Correct answer and approach to transfer for Table B	= P(Z > -1.4) – P(Z > 2.32) = [1 - P(Z > 1.4)] – P(Z > 2.32) = [1 – 0.0808] – 0.0102 = 0.909	= P(Z > -2.41) – P(Z > -0.03) = [1 – P(Z > 2.41)] – [1 – P(Z > 0.03)] = [1 – 0.008] – [1 – 0.488] = 0.48
S1's answer	P (-1.4 < Z < 2.32) ↓ = 0.5667 0.898 = 0.3231	P (-2.41 < Z < -0.03 = 0.5040
Analysis of S1's answer	1 – P(Z > .14) = 1 – 0.4443 = 0.5667 1 – P(Z > 2.32) = 1 – 0.0102 = 0.898 0.898 – 0.5667 = 0.3231	P(Z < -2.41) = 0.008 P(Z < -0.03) = 0.512 0.512 - 0.008 = 0.504
S2's answer	–1.4 < Z < 2.32; 0.0808 < Z < 0.0102	–2.41 < Z < –0.03; 0.0080 < Z < 0.4880
Analysis of S2's answer	P(Z > 1.4) = 0.0808 P(Z > 2.32) = 0.0102	P(Z > 2.41) = 0.0080 P(Z > 0.03) = 0.4880
S7's answer	P (-1.4 < Z < 2.32) = P (Z < 2.32) - P (Z > -1.4) = 1 – P (Z > 2.32) – P (Z > 1.4) = 1 - 0.0102 - 0.0808 = 0.909	P (-2.41 < Z < -0.03) = P(Z < -0.03) – P(Z > -2.41) ~~= 1 – (P > 0.03) – P(Z > 2.4)~~ ~~= 1 – 0.488 – 0.0082~~ ~~= 0.5038~~ => P (Z > 0.03) – P (Z > 2.4) = 0.488 - 0.0082

		= 0.4798
Analysis of S7’s answer	The probability of Z bigger than -1.4 and smaller than 2.32 is the probability of Z smaller than 2.32 minus the probability of Z bigger than -1.4	The probability of Z bigger than -2.41 and smaller than -0.03 is the probability of Z smaller than -2.41 minus the probability of Z bigger than -0.03’

Table 7.6 Comparison of correct answer with Table B and analysis of S1’s, S2’s and S7’s answers in Q1.5-Q1.6

** The third, fourth and fifth rows in S7’s answer in Q1.6 were written but deleted.*

S1 was not familiar with the transferring principles of Table B, so he subtracted two probabilities found inside Table B after incorrect transferring steps in Q1.5 and Q1.6. S2 did not consider the minus sign and simply found probabilities inside Table B, but he did not subtract them. He stopped solving after he found the probabilities corresponding to the mistakenly interpreted positive z scores.

S7 interpreted Q1.5 as the probability of ‘Z being bigger than -1.4 and smaller than 2.32’ and thought it equalled the probability of ‘Z being smaller than 2.32’ minus the probability of ‘Z being bigger than -1.4’. Similarly, she interpreted Q1.6 as the probability of ‘Z being bigger than -2.41 and smaller than -0.03’ and thought it equalled the probability of ‘Z being smaller than -2.41’ minus the probability of ‘Z being bigger than -0.03’. S7’s solving steps can be explained in Figure 7.3.

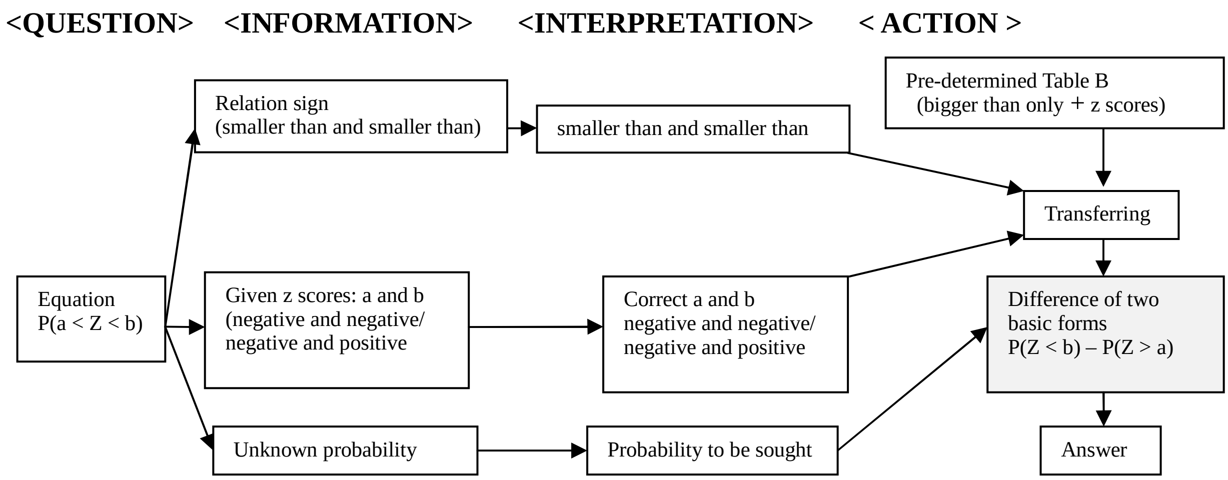


Figure 7.3. S7's solving steps in Q1.5-Q1.6.

She transferred well in the basic questions Q1.1-Q1.4, but she made incorrect transferring steps in Q1.5 and Q1.6. However, some mistakes were cancelled by mathematical luck (Aberdein, 2007), so she still obtained the correct answer on Q1.5 and a close answer on Q1.6.

Q1.5 and Q1.6 aimed to test how students transferred the equation, and half of the participants had difficulties in the transferring steps, especially when dealing with the equation of Z between two values. In the next section, I will discuss how the participants used the tables of normal distribution to find values according to the information given by the questions in different formats which the participants might not notice.

7.3.1.3 Question format

As discussed above, we can see that the requested values (i.e. probabilities) in Q1.1-Q1.6 could be found inside the tables of normal distribution in the direction of 'outside-in' after making the correct transferring steps. However, in the analysis of S4's response to Q1.2, I noticed that she could not recognise the format of the question and of the tables even if they were the same. I also found that some participants did not notice or consider the element of the questions in symbolic form. Therefore, I decided to examine the participants' ability to distinguish the question format in the second interview.

I also tested whether the participants could find values in the tables of normal distribution when the given information was not outside but inside the table of normal distribution. The answer of the question format should be found in the direction of 'inside-out' (see Chapter 5). The new question format allowed me to have a better understanding of the participants' interpretation of the tables of normal distribution and of the questions in symbolic form. I developed Q1.a-Q1.d according to the question in the end of the table:

$P(Z < a) = 0.67$

$P(Z < 0.67) = b$

What is $a + b$?

The question aimed to see whether the participants distinguished what 'a' and 'b' represented respectively. I adopted the two questions and added two more questions with different relation signs. Thus, the four sub-questions were tested at beginning of the second interview as follows:

Q1.a. $P(Z > a) = 0.67$, find 'a';

Q1.b. $P(Z > 0.67) = b$, find 'b';

Q1.c. $P(Z < c) = 0.67$, find 'c'; and

Q1.d. $P(Z < 0.67) = d$, find 'd'.

There is a finding I would like to mention before discussing the participants' solutions to the questions: In the first interview, most participants used only one table of normal distribution, but in the second interview, they switched the table of normal distribution according to the relation sign in the question. Taking S5 and S7 for example:

I: … In the first sub-question, you deleted 0.43. Can you tell me how you found it?

S5: I found it from the table of smaller than (Table A). I haven't used this table for a long time, so I regarded bigger than [i.e. > in the question] as smaller than [which fit the format of Table A).

…

I: Then how did you find that it's wrong?

S5: I looked at them [i.e. Table A and Tale B] and accidentally found that Table B is bigger than and Table A is smaller than.

(S5, 5th Jun 2009).

I: … 'b'?

S7: This is bigger than, so I directly use Table B and get 0.2514.

I: Ok, 'c'?

S7: It is smaller than, so I directly find 0.67 in Table A, and c is 0.44; 'd' is also smaller than, so I can directly use Table A and find 'd' = 0.7486.

(S7, 8th Jun 2009).

After experiencing the exercise questions in the first interview and attending the course for four weeks, the participants seemed to notice the advantage of and difference between Table A and Table B and understand when to use them. This change influenced their solving steps but did not guarantee successful solution of the questions. Whether the participants applied wrong step(s) or did not answer are represented in grey zones in Table 7.7.

	Preferred table in the first interview	Q1.a P(Z > a) = 0.67, find a	Q1.b P(Z > 0.67) = b, find b	Q1.c P(Z < c) = 0.67, find c	Q1.d P(Z < 0.67)= d, find d
S1 used table	A	A	B	A	A
S2 used table	B	A	A	A	
S3 used table	B	A & B	B	A	A
S4 used table	A	B			
S5 used table	B	A (found wrong) then B	A (found wrong) then B	A	A
S6 used table	B	A	B	A	A
S7 used table	B and check with A	B	B	A	A
S8 used table	B	B	B	A	A
S9 used table	B		B		A
S10 used table	B and negative A	A	B	A	A

Table 7.7. Table of normal distribution used by all participants in Q1.a-Q1.d.

**Grey space means that the participant had at least one wrong step or did not answer.*

With consideration of the change of the participants' choice of table of normal distribution and the different question format, the solving steps of Q1.a-Q1.d are different from the solving steps of Q1.1-Q1.6. This can be expanded to Figure 7.4., in which the locations of given numbers and of required answers need to be considered. When a value was given on the left-hand side of the equation, it was a z score, and the participant needed to find the required probability in the direction of outside-in (as was what they did in Q1.1-Q1.6); when a value was given on the right-hand side of equation, it was a probability, and the participant needed to find the required z score in the direction of inside-out.

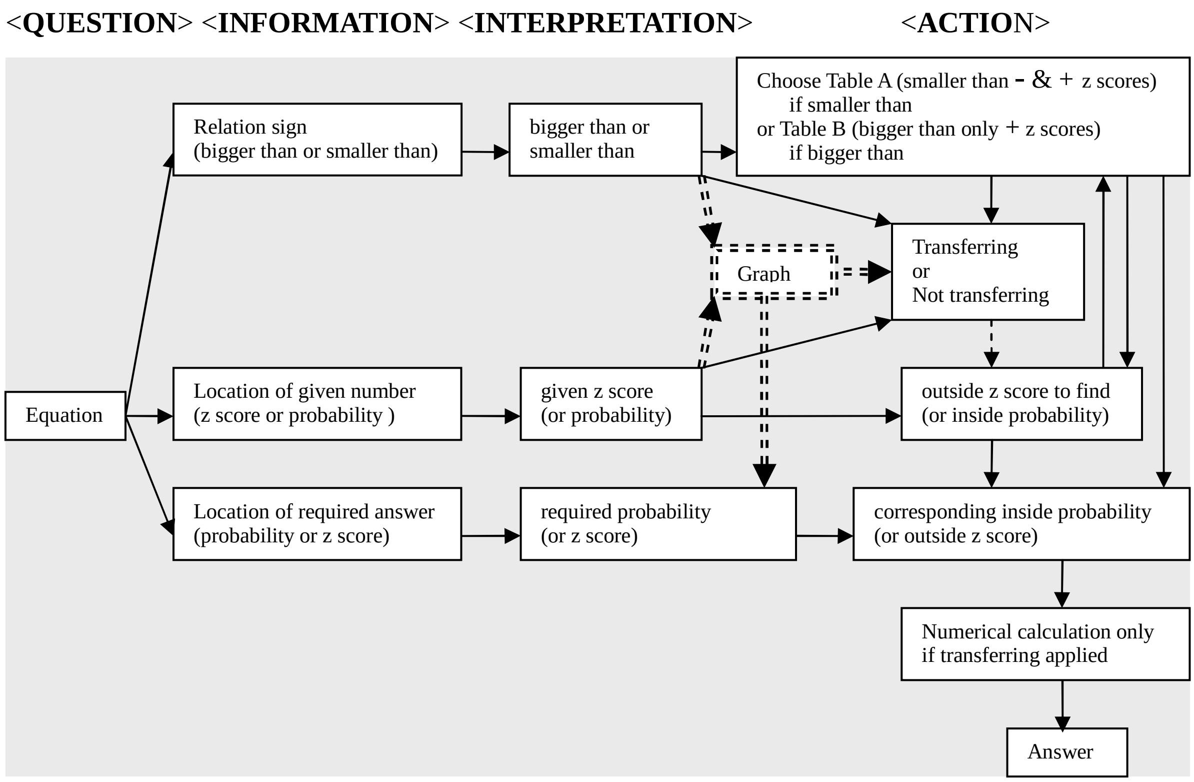


Figure 7.4. Solving processes of using the tables of normal distribution to find the probability according to the given z score or to find the z score according to the given probability.

**The dotted line means transferring may or may not influence the value to be found in the chosen table. This dotted line can be ignored if transferring is not applied.*

*** The 'graph' in broken parallel lines and the three double-dotted lines mean that graphs may or may not be used in the solving processes. They can be ignored if graphs are not applied in the solution.*

****The grey background means that this is the generalised solving process.*

In this section, I will discuss Q1.a first. The 'a' could not be found directly in either Table A or Table B since the format of Q1.a was new and different from that of Table A and that of Table B. Then, I will discuss Q1.c, which was also in a new format, yet its answer could be found outside Table A directly. At the end of this section, I will discuss Q1.b and Q1.d. The format of Q1.b and Q1.d were the same as the format of Q1.4 and Q1.2 respectively and should have been familiar to the participants.

Q1.a-Q1.d represented four kinds of formats, which consist of different given/required information (z score or probability) and relation signs (bigger than or smaller than). I placed Q1.a in the first question in order to test whether the participants recognised the new format. As expected, some students could not distinguish the format of questions. For instance, S4's and S5's answer to Q1.a was the correct answer of Q1.b, and S2's answer to Q1.a was the correct answer of Q1.c.

S4 used Table B and the approach for Q1.b to solve Q1.a because she did not consider the location and meaning of the given/required number, the structure of the question format and the meaning of the question. She routinely solved Q1.a in the 'outside-in' direction to find the given probability 0.67 outside Table B and correspond it to 0.2514 inside Table B as the 'a' (see Figure 7.5.).

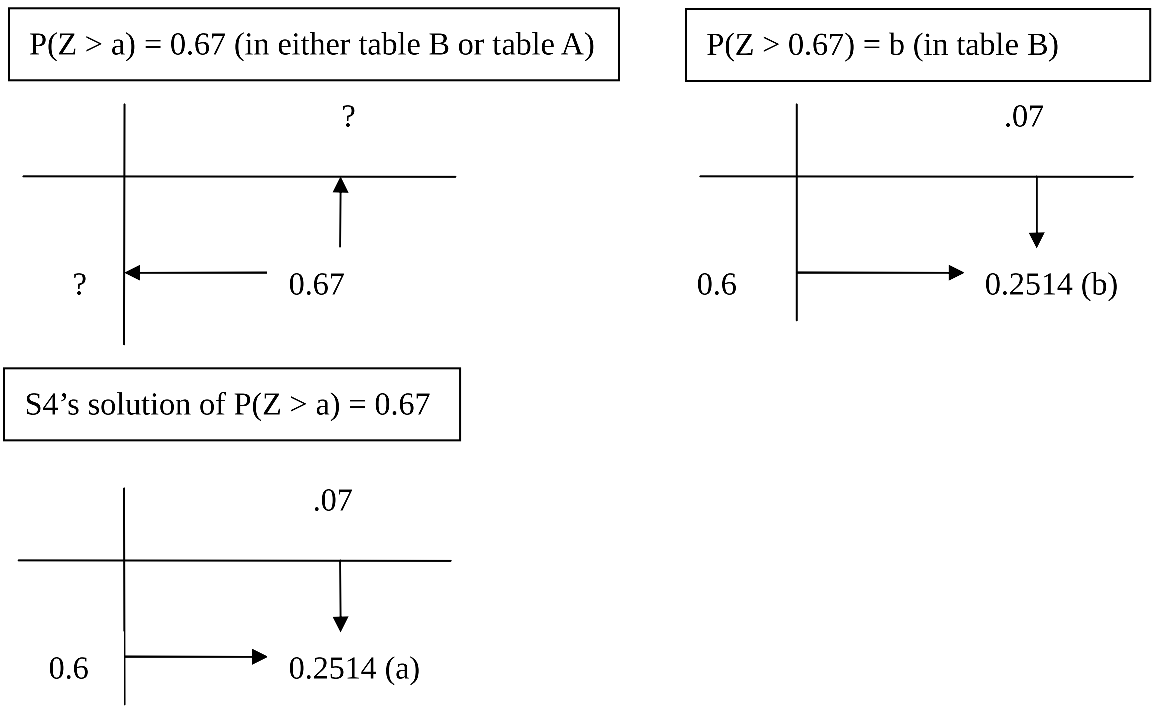


Figure 7.5. Comparison of correct solutions to Q1.a and Q1.b and of S4's solution to Q1.a.

In most arithmetic questions, the required answers or results are on the right-hand side of the equal sign; this was the case with Q1.1-Q1.6 in the first interview. S4 did not recognise the new format of Q1.a, whose answer was on the left-hand side of the equal sign, and applied the typical and familiar method of solving questions in the format of Q1.b, whose required answer was on the right-hand side. S4's solving steps of Q1.a can be expressed in Figure 7.6. She had difficulty distinguishing the formats and ultimately gave up answering Q1.b, Q1.c and Q1.d.

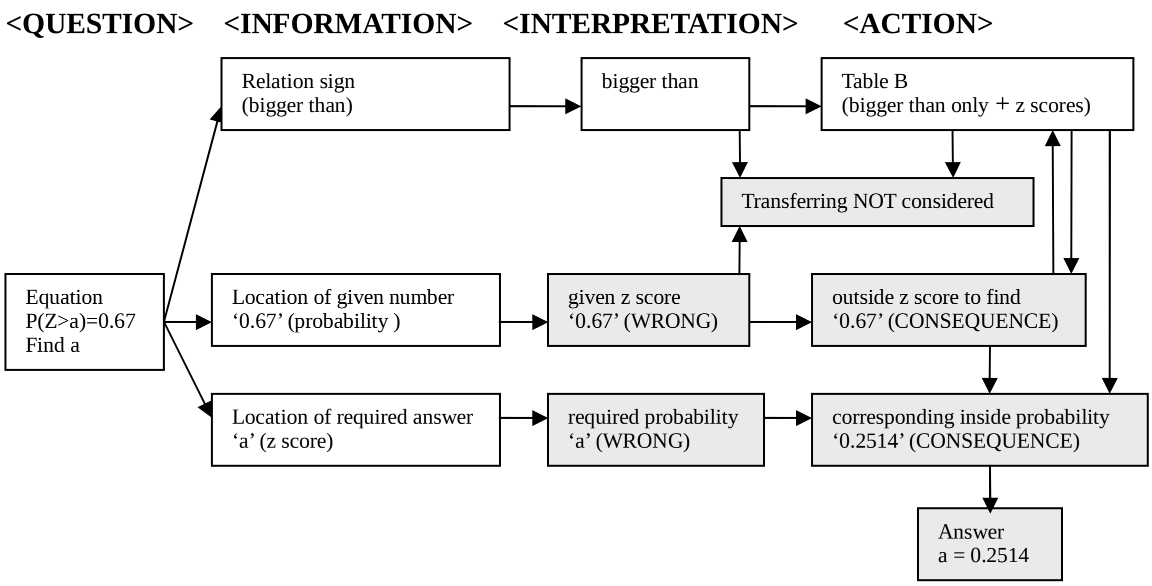


Figure 7.6. S4's and S5's solving process in Q1.a

**The grey zones are (1) steps where mistakes happen or (2) consequences of wrong steps.*

S5 initially found 0.43 (the neighbour of 0.44) outside Table A and 0.7486 inside Table A as the answer to Q1.a and Q1.b respectively because he did not notice the 'bigger than' signs. He soon found his problem when he solved Q1.c, so he revised his answers to Q1.a and Q1.b. He solved Q1.b correctly, but his answer to Q1.a was the same as that of Q1.b. His revised method to solve Q1.a was similar to S4's approach.

In Figure 7.6., we can see that S4 and S5 chose Table B because of the relation sign of ‘bigger than’ and regarded the given number 0.67 on the right-hand side of the equation as a z score. They did not need to transfer the question because they could find 0.67 outside Table B and its corresponding probability .2514 in the direction of outside-in.

S2 only used Table B in the first interview, but he found the given probability 0.67 inside Table A and its corresponding z score 0.44 outside the table. However, his answer of Q1.a was the correct answer of Q1.c, even though he noticed the different relation signs and spent some time to compare the format of Q1.a and Q1.c. Unlike S4 and S5, S2 correctly interpreted the given number on the right-hand side of the equation as a probability and obtained a z score outside Table A in the direction of inside-out as his answer. He noticed the relation signs of ‘bigger than’ in Q1.a but mistakenly believed that Table A should be used to solve questions with relation sign of ‘bigger than’. This can be seen in his solution of Q1.c:

> I: ... The third sub-question is also 0.67, what is c?
> S2: Is it 0.23?
> I: Why 0.23?
> S2: Because it is ‘smaller than’, so 1 – 0.67
> I: 1 – 0.67? It should be 0.33
> S2 laughed: Oh, yes.
> (S2, 1st Jun 2009)

In Q1.a, S2 found the given probability 0.67 inside Table A and its corresponding z score 0.44 outside Table A. However, when he obtained 0.33 by subtracting 0.67 from 1 in Q1.c, he found the 0.33 outside Table A as a z score. The inconsistency of inside-out and outside-in strategies revealed that S2 did not pay enough attention to what kind of value was given and what kind of value was required.

S1, S6 and S10 were the only three participants who had the correct answer to Q1.a. They did not transfer Q1.a to $P(Z < a) = 0.33$ or $P(Z < -a) = 0.67$, but developed a new solving process (see Figure 7.7.):

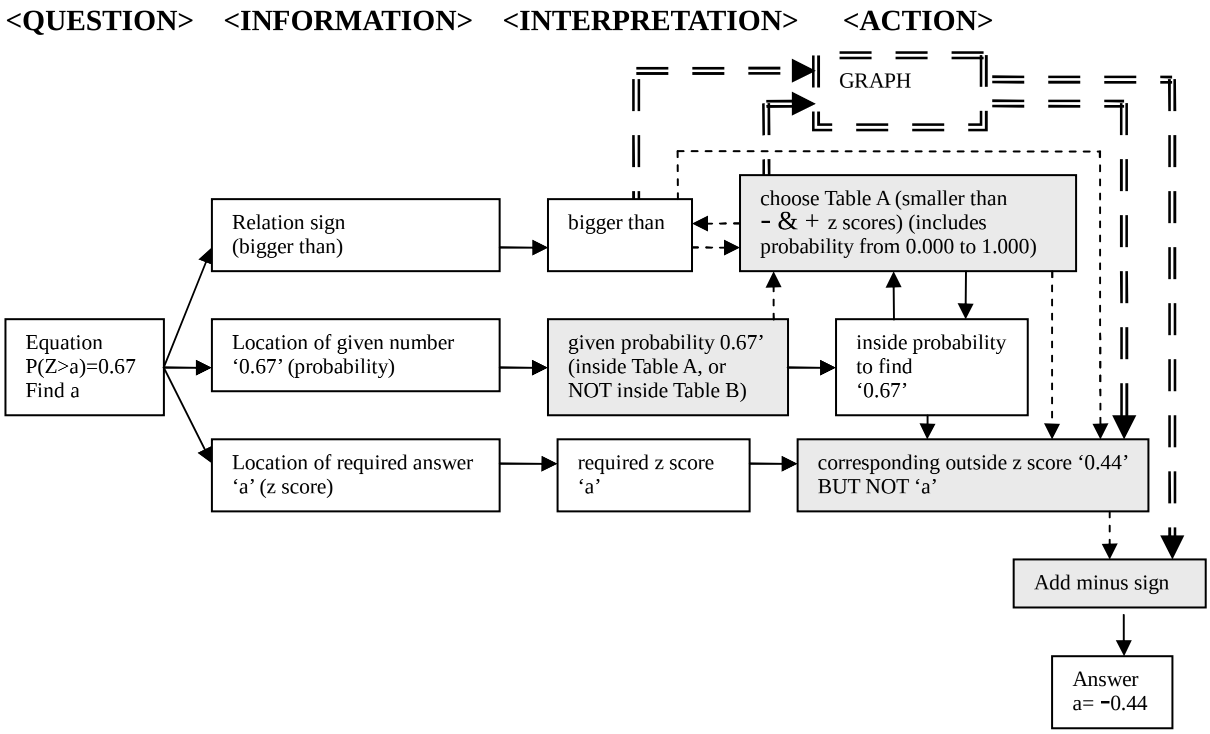


Figure 7.7. S1's, S6's, and S10's solving steps in Q1.a.

** The grey zones and dotted lines represent the series of special thinking approaches of S1 and S6.*

***The ‘graph’ in broken parallel lines and the four double-dotted lines represent how the graph might influence the solving steps. They can be ignored if graphs are not applied.*

These three participants' approach avoided transferring before using Table A. They were clear that Q1.a involved ‘bigger than’ whereas Table A was about ‘smaller than’. They still chose Table A because the given probability 0.67 was inside Table A or not inside Table B. S6 also drew a graph to help consider the relationship of ‘a’ and ‘–a’. These participants found the corresponding z score of 0.44, which implicitly means that P(Z <

0.44) = 0.67, transferred it to $P(Z > -0.44) = 0.67$, and obtained -0.44 as the 'a'. The three participants considered the influence of the relation sign and revised their answers in the last step. The equilibrium of the equation in their solution was maintained.

S1's, S6's and S10's solving steps show that students may develop and apply their own methods to solve new problems. Also, a potential tendency to choose the table of normal distribution was uncovered by comparing these three participants' solving processes in the four sub-questions. The three participants had correct answers in the four sub-questions. In Q1.b, Q1.c and Q1.d, they simply chose the table according to the relation sign of the question and confirmed whether the given number was inside or outside this chosen table. If the given numbers of the three questions could be found inside or outside the chosen table, the table would be 'usable', and transferring steps would not be applied.

This approach could not proceed when solving Q1.a, whose relation sign was 'bigger than' (leading to Table B) but in which the given probability of 0.67 was not inside Table B. None of the three participants transferred the question format to the basic form which would enable them to find the answer in the tables directly. Instead, they found the given probability 0.67 inside Table A first, and then found its corresponding z score 0.44 (i.e. $P(Z < 0.44) = 0.67$). They realised that the meaning of Q1.a was 'the probability of Z being bigger than a was 0.67' (i.e. $P(Z > a) = 0.67$), so they added a minus sign to 0.44 and obtained -0.44 as the answer. However, the equation '$P(Z < 0.44) = P(Z > - 0.44)$' was not written in any of their answer sheets, and they may not have applied the idea of symmetry or recalled the transferring principles in the last step.

The other participants did not correctly solve Q1.a because of different problems. S3 solved Q1.b, Q1.c and Q1.d correctly and only had a problem in Q1.a. Unlike S1, S6 and S10, S3 tried three times to transfer Q1.a to another format before choosing the table of normal distribution. Initially, she wrote:

$$P(Z > a) = 0.67$$
$$1 - P(Z > a) = 1 - 0.67$$

She gave up this solution quickly and retried two times, but she mis-transferred the equations and obtained the wrong answer:

$$P(Z < -a) = 1 - 0.67 = 0.33$$
$$-a = -0.44$$
$$a = 0.44$$

$$P(Z > a) = 0.67$$
$$1 - P(Z > a) = 0.67$$
$$P(Z > a) = 0.33$$
$$a = 0.44$$

S3's mis-transferring steps generated incorrect information and resulted in the two trials of using Table A and Table B respectively. Unfortunately, she still got the wrong answer.

S8 used Table B to solve Q1.a in terms of the 'bigger than' sign, but she mis-transferred the equation and obtained 1 – 0.67 = 0.33. She also looked for 0.33 inside Table B, and her solution of Q1.a was similar to S3's.

S7 also solved Q1.b, Q1.c and Q1.d without difficulty, and she had the answer 0.44 for Q1.a at the beginning, but she gave up this answer when she noticed that "it involves bigger than and smaller than" (S7, 8th Jun 09). In her second trial, she mis-interpreted the numbers inside the table, and this made her feel that the probability 0.67 was between .0681 and .0668 and that transferring was not needed

In the analysis of the participants' solutions in Q1.a, we can see that this question was difficult for the participants, so only three of them obtained the correct answer. Also, their difficulties occurred in different aspects, such as not considering the relation sign, regarding the given probability as z score or making the wrong transferring steps. Below I discuss their solutions to Q1.c, which was also in a new question format.

Q1.c was in a new format that the participants were not familiar with, but it seemed to be easier than Q1.a because its answer could be found in Table A directly. S4 and S9 did not answer Q1.c, and the other eight participants explained that they chose Table A to solve Q1.c because the question was about 'smaller than'. Six of them (except for S2 and S8) had correct solving steps, as shown in Figure 7.8.

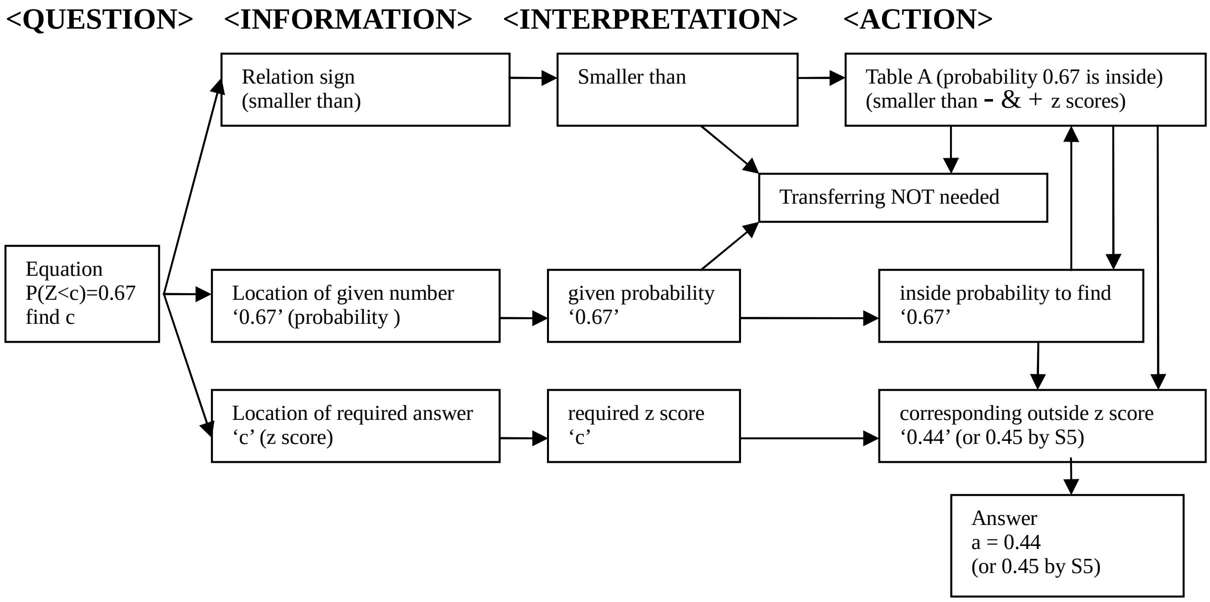


Figure 7.8. Six participants' solving steps in Q1.c.

When S2 solved Q1.c, he noticed that the relation sign of the question was 'smaller than', which matched the format of Table A. He obtained 0.33 by subtracting 0.67 from 1, but he found 0.33 outside Table A and corresponded it to 0.6293 inside Table A as the answer of 'c'.

S2's subtraction made sense since P(Z < c) = 0.67 can be transferred to 1 – P(Z < c) = 1 – 0.67 and P(Z > c) = 0.33 and c could be found outside Table B. However, S2 believed that Table A should be used when the relation sign was bigger than, and this also explained why he used Table A in Q1.a. Thus, he applied an unnecessary transferring in Q1.c and found the (probability) value outside Table A. His solving process of Q1.c was completely different from that of Q1.a. S8 also solved Q1.c in a similar way (see Figure 7.9.):

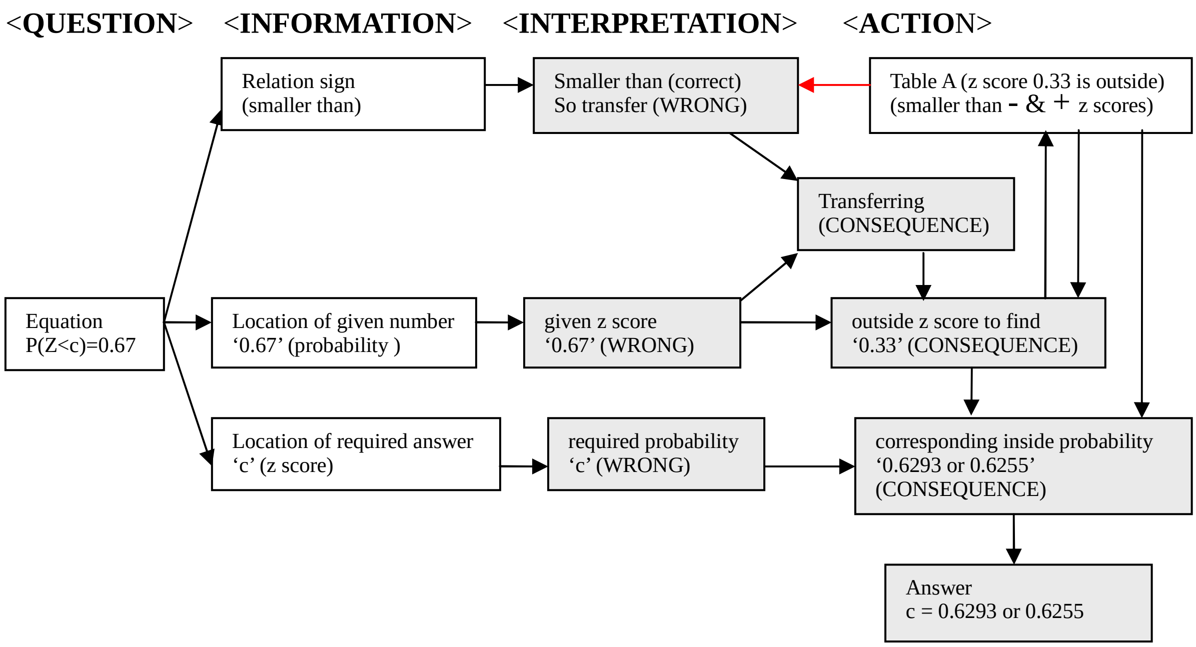


Figure 7.9. S2's and S8's solving processes of Q1.c.

** Grey zones are (1) steps where mistakes happen or (2) consequences of wrong steps*

Below I will discuss Q1.b and Q1.d together because the formats of Q1.b and Q1.d were the same as the formats of Q1.4 and Q1.2 respectively, and the formats had been tested in the first interview and were thus not new to the participants. How the participants solved questions in the two formats in the first interview has been discussed in Section 7.3.1.1.

In the second interview, Q1.b and Q1.d were not problematic for most participants, and six participants (S1, S3, S6, S7, S8 and S10) used correct solving strategies in both questions (see Table 7.8). S5 used the correct approach in Q1.d and actively corrected his wrong approach in Q1.b after he compared the format of Q1.b and Q1.d. Therefore, Q1.b and Q1.d were problematic only for S2, S4 and S9.

Question number	Q1.2	Q1.4	Q1.b	Q1.d
Question	P(Z < 1.22) = ?	P(Z > 2.05) = ?	P(Z > 0.67) = b, find 'b'	P(Z < 0.67) = d, find 'd'
Correct answer and approaches	**= 1 – P(Z > 1.22) = 1 – 0.1112 = 0.8888 (with Table B) [or =0.8888 with Table A, not used]**	**= 1 – P(Z < 2.05) = 1 – 0.9798 = 0.0202 (with Table A) or = P(Z < -2.05) = 0.0202 (with Table A) or = 0.0202 (with Table B)**	**P(Z > 0.67) = 0.2514 (with Table B) [or P(Z > 0.67) = 1 - P(Z < 0.67) = 1 – 0.7486 = 0.2514 (with Table A) or P(Z > 0.67) = P(Z < -0.67) = 0.2514 (with Table A), not used]**	**P(Z < 0.67) = 0.7486 (with Table A) [or P(Z < 0.67) = 1 - P(Z > 0.67) = 1 – 0.2514 = 0.7486 (with Table B), not used]**
S2's answer	< 1.22; 0.1112 (with Table B)	> 2.05; 0.0202 (with Table B)		
S4's answer		0.9798 (with Table A)		
S9's answer	= P($\frac{x-\mu}{\sigma_x}$<1.22) = P(x > 1.22) = 0.1112(with Table B)	= P($x-\mu/\sigma$ >2.05) = P(x > 2.05) = 0.0202 (with Table B)	0.2514	-0.2514
Number of participants who used correct approach	6	9	7+1 (S5)	7

Table 7.8. Three participants' difficulties in questions of two formats.

S2 and S4 did not answer Q1.b and Q1.d. The other eight participants (including S5, who actively revised his wrong strategy) used the correct strategy and had correct answers to Q1.b. Their solving processes are displayed in Figure 7.10.

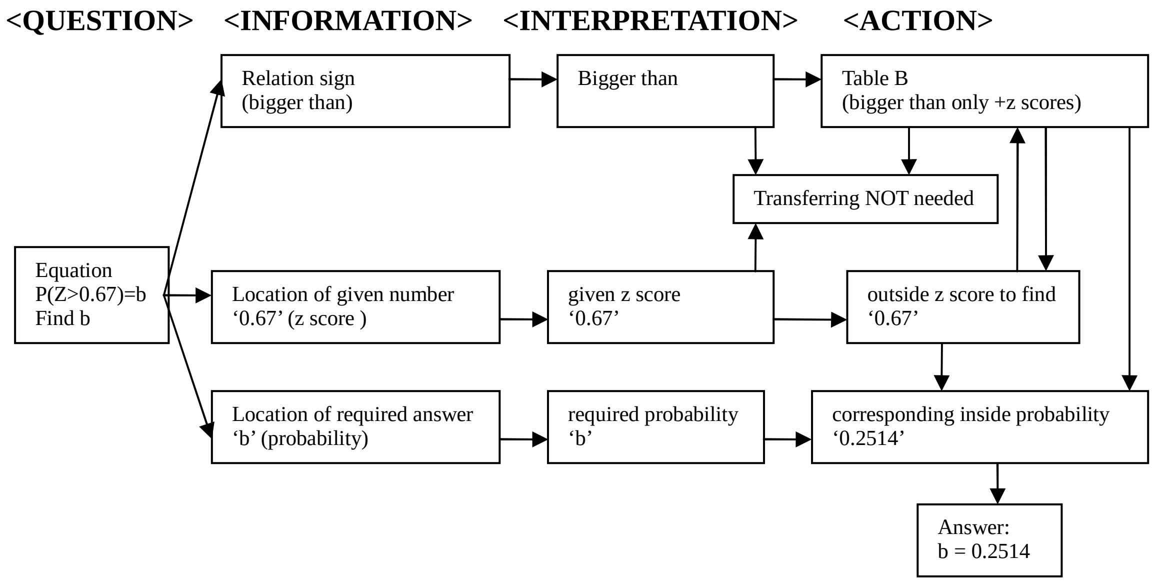


Figure 7.10. Eight participants' solving steps in Q1.b.

In Figure 7.10, we can see that the eight participants chose Table B because of the relation sign of ‘bigger than’. They correctly interpreted the given number 0.67 as z score and the required number as probability. Moreover, the z score 0.67 could be found outside Table B, so transferring was not needed before using Table B. Therefore, the participants simply found the corresponding probability 0.2514 inside Table B as the answer to Q1.b. There was no further calculation needed since transferring was not applied.

Seven of the eight participants (except for S9) who correctly solved Q1.b also applied the correct strategy in Q1.d (see Figure 7.11.). They chose Table A because of the relation sign of ‘smaller than’. The given number 0.67 was still a z score and the required number was a probability. The z score 0.67 also could be found outside Table A, so transferring was not needed before using Table A. Thus, the participants found the corresponding probability 0.7486 inside Table A as the answer to Q1.d.

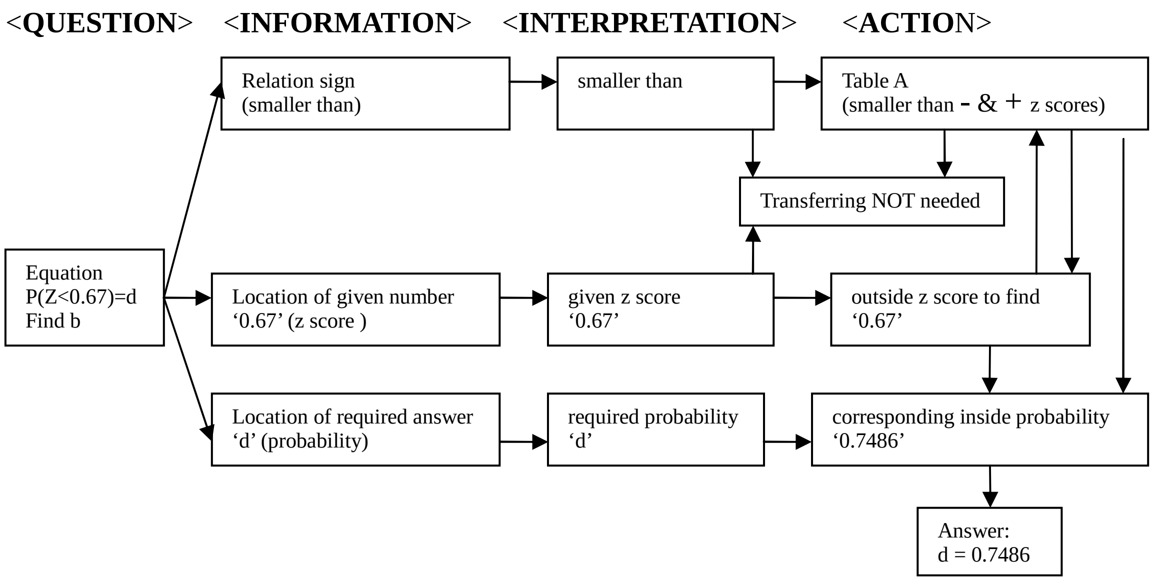


Figure 7.11. Seven participants' solving steps in Q1.d.

S9 did not notice the existence of the second page of Table A, so he only used the first page of Table A (i.e. negative z scores) to find P(Z < -0.67) = 0.2514 and transferred it to P(Z < 0.67) = -0.2514. His solving steps of Q1.d can be seen in Figure 7.12.

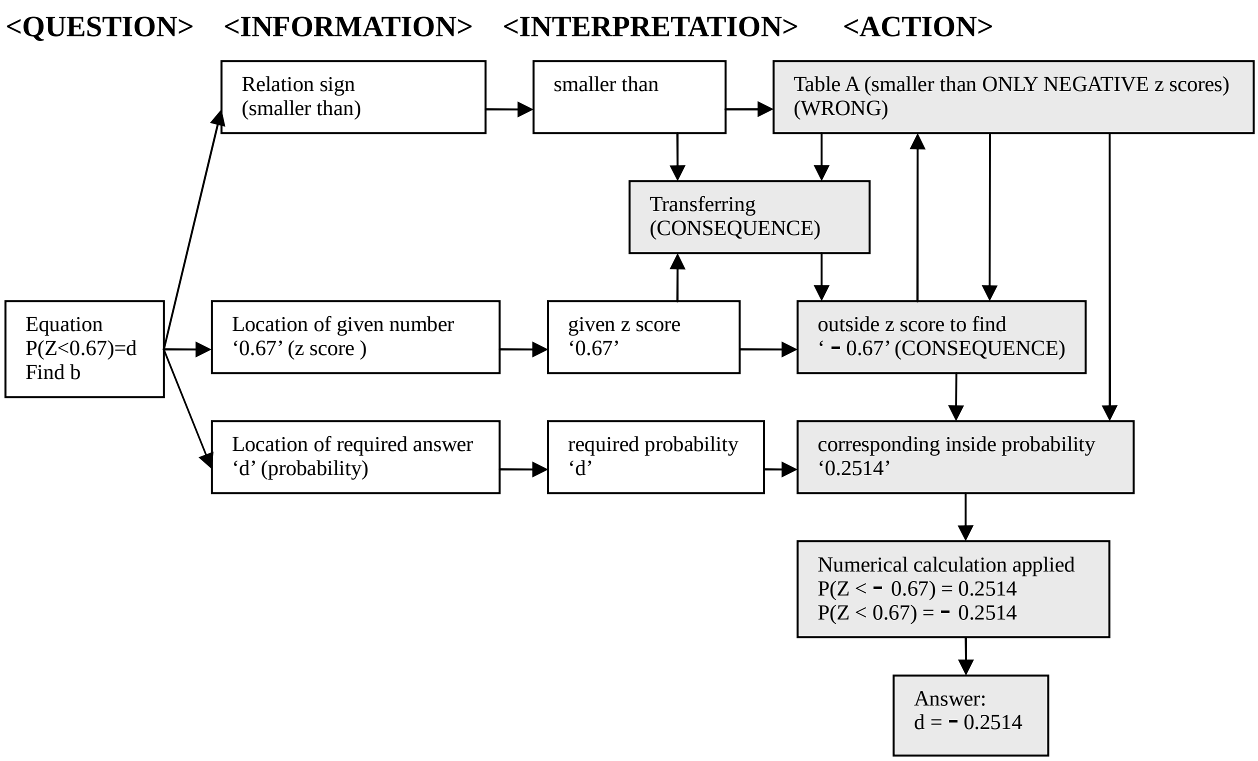


Figure 7.12. S9's solving steps in Q1.d

Through Figure 7.12, we can see that the absence of the positive part of Table A prevented S9 from directly finding the 0.67 outside Table A and led him to use -0.67. He appeared to transfer P(Z < -0.67) = 0.2514 to P(Z < 0.67) = -0.2514. In his solutions in Q1.d and Q1.1 (see Table 7.9.), S9 had difficulty transferring the format of 'smaller than a negative z score'. However, after being told the existence of the positive part of Table A, S9 found the correct answer quickly. He explained that he had overlooked the positive part because of his unfamiliarity with Table A.

Question number	Q1.1	Q1.d
Question	P(Z < -0.03) = ?	P(Z < 0.67) = d, find d
Correct answer and approaches	**= P(Z > 0.03)** **= 0.4880 (with Table B)** **[or = 0.4880 (with Table A, not used)]**	**P(Z < 0.67) = 0.7486 (with Table A)** **[or P(Z < 0.67)** **= 1 - P(Z > 0.67)** **= 1 – 0.0202** **= 0.7486 (with Table B), not used]**
S9's answer	= P($\frac{x-\mu}{\sigma_x}$ < -0.03) = P(x < -0.03) = 1 - P(x > 0.03) = 1 – 0.4880 = 0.5120 (with Table B)	-0.2514 (with negative part of Table A)
Analysis of S9's answer		P(Z < -0.67) = 0.2514 P(Z < 0.67) = -0.2514

Table 7.9. S9's problem in dealing with two questions of 'smaller than'.

In Section 7.3.1, I have discussed how the questions in symbolic form caused the participants' difficulties. In the next two sections, I will discuss two more factors which caused them difficulties.

7.3.2 Using the table(s): Confusion of what is inside/outside the table(s)

No matter if the students could interpret the question correctly or not, they might find the given number in a wrong place of the chosen table. For example, they might find a given z score inside or a given probability outside the table of normal distribution (i.e. mis-interpreting). In other words, they might not use the table in the correct direction of ‘outside-in’ or ‘inside-out’, especially when the value of given probability was included and could be found directly outside the table, or when the value of given z score was included and could be found inside the table. For example, S4 and S5 found the probability outside Table B in Q1.a, and S2 and S8 found the probability outside Table A in Q1.c (see Figure 7.13.). Such problems happened more often when the participants could not recognise the question formats and got confused about which kind of information (i.e. z score or probability) was given/required.

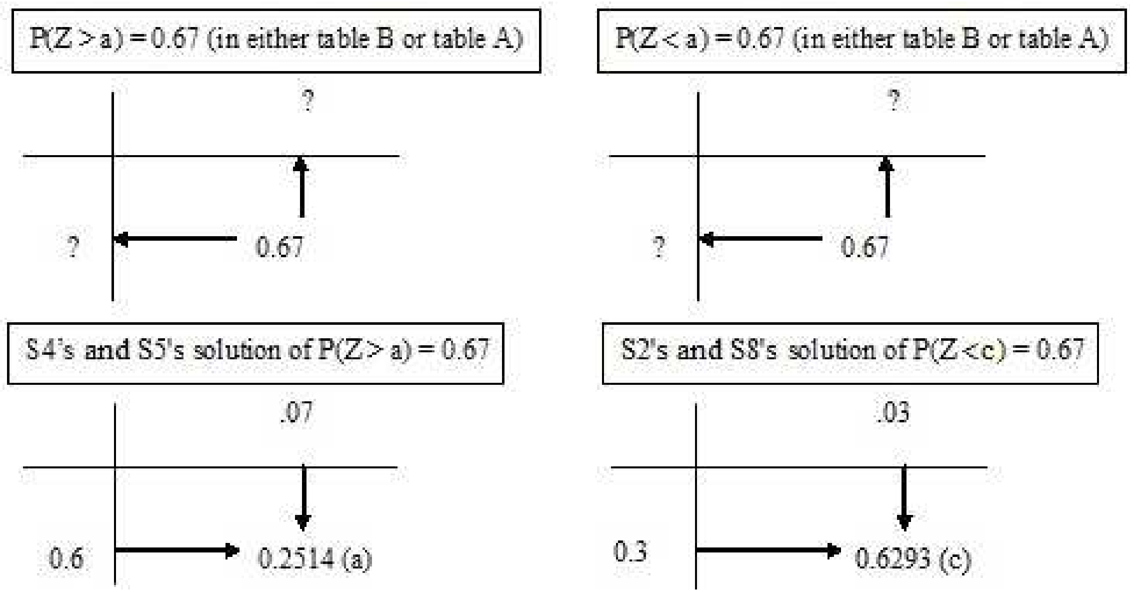


Figure 7.13. Comparison of correct solution and S4’s and S5’s solution in Q1.a, and comparison of correct solution and S2’s and S8’s solution in Q1.c

7.3.3 When the answer was not visible in the table: Negative answer

By examining the participants’ solving processes and answers to Q1.1-Q1.4 and Q1.a-Q1.d, it is evident that most participants had difficulty finding the answer which could not be found directly in the table without transferring, especially when the answer(s) was/were negative. This situation was also found in Q3.2, which tested the use of the table of t distribution (see Section 8.3).

In the eight questions of using the tables of normal distribution to find a value (i.e. Q1.1-Q1.4 and Q1.a-Q1.d), more than half of the participants (normally six or seven of them) applied correct methods in the questions, with the exception of Q1.a, which appeared to be difficult for the participants, as only three of them had the correct solving process. Q1.a had a special feature in that the unknown answer was negative and could not be found directly in both tables of normal distribution. The participants might not consider the unknown value as negative, so six of them obtained positive answers.

7.4 Summary

In the beginning of this chapter, I separated the exercise questions in the clinical interviews into four categories. I also categorised seven symptoms of participants' difficulties in their solutions. I then analysed the participants' solutions according to the framework of solving steps and determined several factors causing their difficulties when using the tables of normal distribution to find a value. In the first interview, I found that most participants only used one preferred table of normal distribution and had difficulty when the questions contained a negative number or when the answer was not visible in the table and could not be found directly in the table of normal distribution without transferring. This difficulty became more obvious when they attempted to solve Q1.a, which was in a different format and had a negative answer, even if they chose the table of normal distribution according to the relation sign in the second interview.

In the next chapter, I will discuss the participants' performance in using the table of t distribution to look for t score according to the given probability and degree of freedom.

Because t distribution was not emphasised by the teacher, there were only two questions about t distribution in the interviews.

CHAPTER 8

DATA ANALYSIS II

APPLICATION OF THE TABLE OF T DISTRIBUTION

The table of t distribution is important since the population standard deviations σ are always unknown. The t distribution is always used for small samples. However, unlike the unique standard normal distribution, there are many t distributions. Each t distribution is decided by the degree of freedom, which equals the sample size minus one (Wonnacott and Wonnacott, 1990). I will not discuss detailed theory of t distribution in this study. Rather, I will focus on how the participants used the table of t distribution to find the t score according to the given probability and degree of freedom (or sample size). Such a format was the only format of t distribution used in the participants' class (as the key step to find confidence interval or to find critical region in hypothesis testing).

8.1 Participants' performance in two questions

Following Q1.1-Q1.6, Q3.1 and Q3.2 asked the participants to use the attached table of t distribution (see Appendix 2.3.) to find the values:

Q3.1 Look for $t_{0.005}(9)$; and

Q3.2 Look for $t_{0.99}(23)$

The questions, answers and the participants' performance are listed in Table 8.1. We can see that most participants used the correct method in Q3.1, but most of them used an incorrect method in Q3.2 (for the details, see Appendix 8.1.).

Question number	Q3.1	Q3.2
Question	Look for $t_{0.005}(9)$	Look for $t_{0.99}(23)$
Correct answer	3.25	-2.5
Number of participants with correct approach	9	2 (S1, S7)
Number of participants with incorrect approach	0	7
Number of participants who had no answer	1 (S9)	1 (S9)

Table 8.1. Questions, answers and the participants' performance in Q3.1 and Q3.2.

8.2 Basic format of t distribution

The first question of t distribution (Q3.1) was not problematic for most participants since the given probability and the v (degree of freedom) were included and could be found in the table of t distribution. The only exceptions were S9, who could not use the table of t distribution, and S7, who subtracted 1 from the given v before using the table of t distribution. S7 explained that

> I did not know that I should directly look at the v given in the question. I remember that there was a 'minus one' in somewhere to get v (S7, 4th May 09).

The solving steps of Q3.1 are much different from those of Q1.1-Q1.6 and Q1.a-Q1.d, and eight participants successfully solved Q3.1 in this way (see Figure 8.1.):

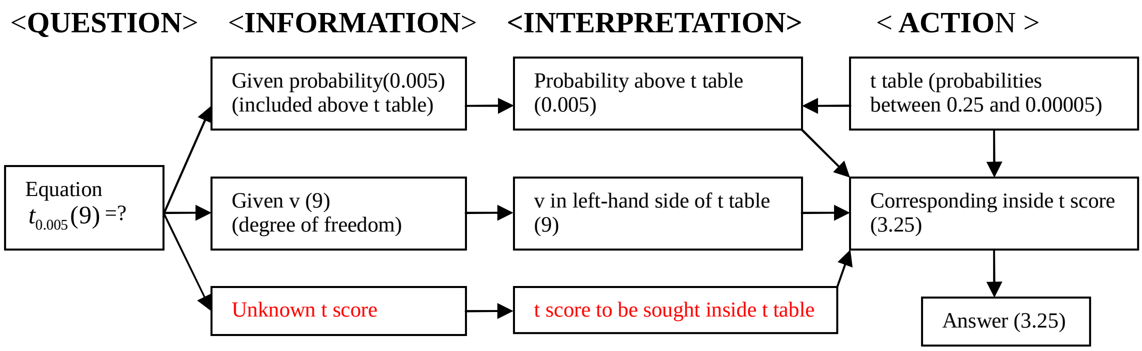


Figure 8.1. Eight participants' (excluding S7 and S9) solving steps in Q3.1.

S7's solving strategy was similar to Figure 8.1., but she used a wrong degree of freedom, so her answer was incorrect. Although Q3.1 was successfully solved by most participants, Q3.2—whose negative answer was not included in the table of t distribution—appeared to be one of the most challenging questions for them in the entire study.

8.3 Advanced question with negative t score

As discussed in Section 7.3.1, when solving Q1.1-Q1.6, some participants did not pay attention to whether the given values were positive or negative and appeared to treat all numbers as positive. This situation became more obvious when the unknown answer was negative. Q3.2, which has a negative answer that cannot be found directly in the table of t distribution, was apparently difficult for the participants. Only two participants realised that the answer to Q3.2 was a negative value, and another two participants obtained negative answers by calculation by applying incorrect solving strategies. Three other participants had positive answers, and the other three participants did not know where or how the answer could be found.

Only S1 and S7 applied the correct approach to solve Q3.2, but S7 made the same mistake as she had made in Q3.1. Three participants, including S3, who drew a graph (see Figure 8.2.), did not have the answer. S3's graph did not match the condition of the question and was therefore of no help in solving the question. I will discuss the participants' graphs in Chapter 11.

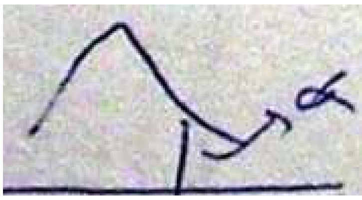


Figure 8.2. S3's graph in Q3.2.

The other five participants noticed that the given probability 0.99 was not included above the table of t distribution. Therefore, they obtained 0.01 by subtracting 0.99 from 1 because they found that there was 0.01 on top of the table of t distribution. Their solving processes, answers and an analysis of their answers are listed in Table 8.2.

	Correct	S4	S5	S6	S8	S10
Answer	$t_{0.99}(23)$ $=-t_{0.01}(23)$ $=-2.5$	$t_{0.99}(23)$ $=1.319$	$t_{0.99}(23)$ $=1-0.01$ $=1-2.500$	$t_{0.99}(23)$ $=31.821-2.5$	$t_{0.99}(23)=-1.500$ $1-0.99=0.01$ $1-2.5=-1.5$	$t_{0.99}(23)$ $=31.821-2.5-2.5$
Analysis of answer		$t_{0.1}(23)$ $=1.319$		$t_{0.01}(1)=31.821$ $t_{0.01}(23)=2.5$		$t_{0.01}(1)=31.821$ $t_{0.01}(23)=2.5$

Table 8.2. Analysis of five participants' answers to Q3.2.

Four of the five participants found the t score $t_{0.01}(23)$ = 2.5 corresponding to the probability 0.01 and v 23, but their problems occurred in the consequences of transferring. Both S5 and S8 obtained 0.01 by subtracting the given probability 0.99 from 1, and then found 0.01 outside the table of t distribution. When they found the t score 2.5 inside the table corresponding to 0.01, they subtracted 2.5 from 1 without considering the meaning. In other words, they imitated their previous method of dealing with probability and applied it again to transfer the found t score. Interestingly, both S6 and S10 created $t_{0.01}(1)$ rather than $t_1(23)$ in their transferring processes. Then S6 subtracted $t_{0.01}(23)$ from $t_{0.01}(1)$ and S10 subtracted $t_{0.01}(23)$ from $t_{0.01}(1)$ two times. S6 used the graph above the table of t distribution, but the graph did not help. This will be further discussed in Chapter 11.

The four participants' responses showed that the meanings of the elements in the question and in the t table were not considered in their solving procedures. For example, they subtracted the given 0.99 from '1' and got 0.01, but they seemed not to realise why '1' should be used and what '1' meant. Moreover, they did not know how the introduced '1' influenced the found t score. Thus, the '1' remained for further transferring after they found the t score corresponding to the generated 0.01. The participants' solving strategy seemed to be influenced by their approach to solving previous questions of using tables of normal distribution, such as $P(Z < 1.22) = 1 - P(Z > 1.22)$ and $P(Z > -3.1) = 1 - P(Z > 3.1)$.

S4 also obtained 0.01, but she mistakenly interpreted .1 above the t table as 0.01, so she found a t score of 1.319. That she did not transfer afterwards matched her solving steps in Q1.1-Q1.6, in which no transferring step was applied. The solving steps of participants who had answers for Q3.2 are included in Figure 8.3.

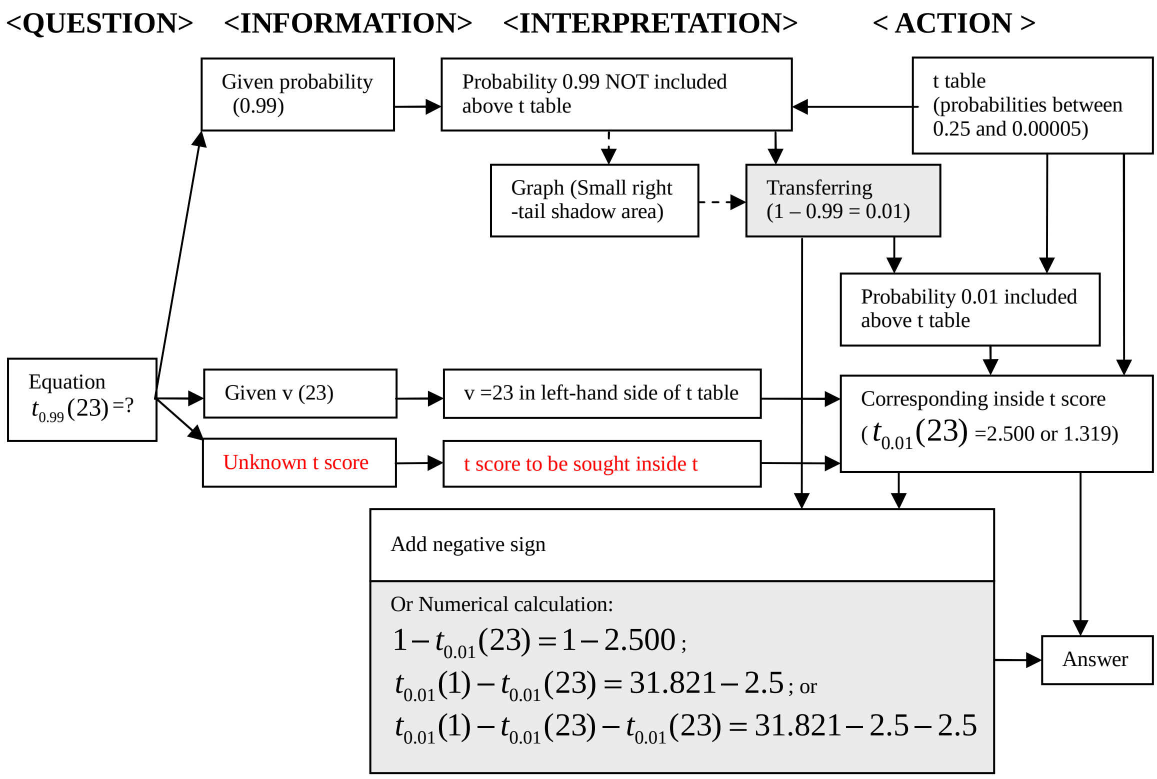


Figure 8.3. Five participants' solving steps in Q3.2.

**The dotted lines mean that participants may not take advantage of graphs when the given probability is not included above the t table and transferring is needed.*

***The grey zones mean that the transferring caused students to apply incorrect numerical calculations.*

8.4 The participants' difficulties and causes in Q3.1 and Q3.2

The participants' difficulties in Q3.1 and Q3.2 and the causes of these difficulties are summarised in Table 8.3. We can see that participants' difficulties were mainly caused by the unfamiliar question format of a negative answer which cannot be found in the table of t distribution directly.

<table>
<tr><td colspan="3">Stage
(Careless) mistakes & Causes</td><td>Question</td><td>Information</td><td>Interpretation</td><td>Action</td></tr>
<tr><td rowspan="6">(Careless) mistakes</td><td colspan="2">Mis-reading</td><td>-</td><td>-</td><td>0</td><td>-</td></tr>
<tr><td colspan="2">Mis-transcribing</td><td>-</td><td>-</td><td>-</td><td>0</td></tr>
<tr><td colspan="2">Mis-transferring</td><td>-</td><td>-</td><td>-</td><td>5</td></tr>
<tr><td colspan="2">Mis-targeting</td><td>-</td><td>-</td><td>-</td><td>0</td></tr>
<tr><td colspan="2">Mis-interpreting</td><td>-</td><td>-</td><td>-</td><td>3</td></tr>
<tr><td colspan="2">Mis-calculation</td><td>-</td><td>-</td><td>-</td><td>0</td></tr>
<tr><td rowspan="3">Causes</td><td>Reading the question in symbolic form</td><td>Question format</td><td>-</td><td>-</td><td>10</td><td>7</td></tr>
<tr><td colspan="2">Using the table
(The degree of freedom)</td><td>-</td><td>-</td><td>-</td><td>2</td></tr>
<tr><td colspan="2">When the answer was not in the table (Negative answer)</td><td>-</td><td>-</td><td>8</td><td>5</td></tr>
</table>

Table 8.3. Frequencies of symptoms and causes in participants' solutions in Q3.1 and Q3.2.

** Numbers in the table mean how many times such situations happened; repeated symptoms in the same question were counted multiple times '-' means inappropriate.*

***The counting of the numbers is explained in Appendix 7.3.*

In the analysis of Q3.1 and Q3.2, we can see that Q3.1, which was in the basic format of the table of t distribution, was correctly solved by most participants, and that Q3.2, whose negative answer could not be found directly in the table of t distribution, was challenging to most of them. Symbolic manipulation enables them to solve some questions, but it may have constraints in some other questions.

8.5 Summary

In this chapter, I explained how the participants solved the two questions of using the table of t distribution to find values. Most of them had difficulty in Q3.2 because the given value was not included in the table and the participants did not consider the answer as a negative. Some participants totally relied on symbolic manipulation to solve the question without considering the meaning of the question and the symbols. Most of them did not use graphs, and only two participants used graphical representation in Q3.2 (see Chapter 11), but their graphs did not help in solving the question. Since the t distribution was not emphasised in the participants' class, I did not have other questions about the t distribution and its table.

In Chapter 9 and Chapter 10, I will discuss how the participants found the frequently used z score of 90%, 95% and 99% confidence intervals, and how they found the critical value/region in hypothesis testing. Then I will discuss the influence of visualisation in Chapter 11.

CHAPTER 9

DATA ANALYSIS III

Z SCORE OF CONFIDENCE INTERVAL

In the previous two chapters, I discussed the participants' use of tables of normal distribution and their use of the table of t distribution. In this chapter, I will discuss how the participants learned one of the most important ideas in statistical inference: confidence interval. First, I will introduce the definition and formula of confidence interval and explain how to find the three frequently used z scores of 90%, 95% and 99% confidence intervals in the tables of normal distribution. Then, I will examine how the participants interpreted confidence interval and whether they were able to find the three commonly used z scores in the tables of normal distribution by themselves, even if many of the participants might have memorised the values. The questions were tested at the end of the first interview as Q15.

9.1 Definition and formula of confidence interval

Confidence interval is an early and practical application of z scores and is the result of interval estimation.

> $\overline{X}$ is a good estimator of μ for populations that are approximately normal. Although on average $\overline{X}$ is on target, the specific sample mean $\overline{X}$ that we happen to observe is almost certain to be a bit high or a bit low. Accordingly, if we want to be reasonably confident that our inference is correct, we cannot claim that μ is

precisely equal to the observed $\overline{X}$. Instead, we must construct an interval estimate or confidence interval of this form: " $\mu = \overline{X} \pm$ sampling error" (Wonnacott & Wonnacott, 1990, p. 254).

9.2 Z scores of 90%, 95% and 99% confidence interval

The width allowed for the sampling error depends on how much $\overline{X}$ fluctuates and can be decided by how confident the user wishes the interval estimate to be correct. Frequent choices are 90%, 95% and 99%, and the most frequent choice is 95% confidence interval, which means a correct interval 19 times out of 20. Therefore, the selected area below the normal curve in the "middle chunk" will enclose a 95% probability and leave 2.5% probability excluded in each tail. This leads to z scores 1.96 and -1.96, i.e. P(Z > 1.96) = 0.025 and P (Z < -1.96) = 0.025. Therefore, we can get

$$P(-1.96 < Z < 1.96) = 0.95$$

which is the standardised equation of

$P(X - 1.96SE < \mu < X + 1.96SE) = 95\%$ (SE is the abbreviation of standard error $= \frac{\sigma}{\sqrt{n}}$).

It is important that μ is not a variable but a constant, whereas the interval $(\overline{X} - 1.96SE, \overline{X} + 1.96SE)$ varies (Wonnacott & Wonnacott, 1990).

In most statistics courses, 1.645 and -1.645 of 90% confidence interval and 2.575 and -2.575 of 99% confidence interval are also frequently used. Therefore, there are three useful formulas:

90% confidence interval

$$\mu = \overline{X} \pm z_{0.5}\frac{\sigma}{\sqrt{n}} \;=\; \overline{X} \pm 1.645\frac{\sigma}{\sqrt{n}} \;\text{ or }\; (\overline{X} - 1.645\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.645\frac{\sigma}{\sqrt{n}});$$

95% confidence interval

$$\mu = \overline{X} \pm z_{0.25}\frac{\sigma}{\sqrt{n}} \;=\; \overline{X} \pm 1.96\frac{\sigma}{\sqrt{n}} \;\text{ or }\; (\overline{X} - 1.96\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}); \text{ and}$$

99% confidence interval

$$\mu = \overline{X} \pm z_{0.05}\frac{\sigma}{\sqrt{n}} \;=\; \overline{X} \pm 2.575\frac{\sigma}{\sqrt{n}} \;\text{ or }\; (\overline{X} - 2.575\frac{\sigma}{\sqrt{n}}, \overline{X} + 2.575\frac{\sigma}{\sqrt{n}})$$

Therefore, 99% confidence interval gives the users more confidence than 95% and 90%. However, the estimation is not more exact, but wider and vaguer. Wonnacott and Wonnacott (1990) explain that “the more certain we want to be about a statement, the more vague we must make it” (1990, p.259).

Before I started collecting data, the teacher explained the idea of the z scores of 90%, 95% and 99% confidence intervals by drawing graphs such as Figure 9.1.

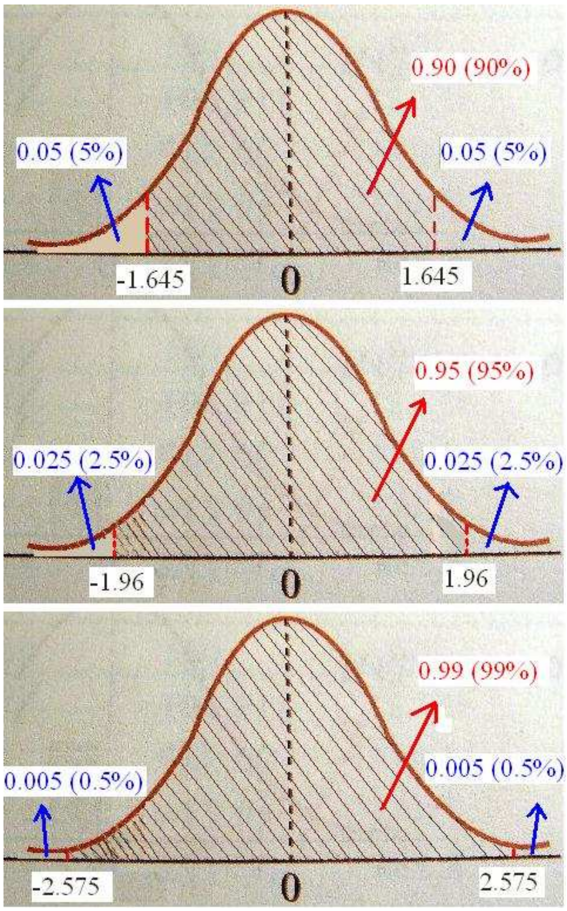


Figure 9.1. Graphical representation of z scores and probabilities of 90%, 95% and 99% confidence intervals.

Although the relationship between the probability (i.e. 90%, 95% or 99%) and the z score was explained with graphs in class, the participants were asked by the teacher to 'memorise' the three frequently used z scores 1.645, 1.96 and 2.575 of 90%, 95%, and 99% confidence intervals respectively. Thus, whether and how the participants realised the relationship between the z scores and probabilities should be examined. In Q15, all the participants were asked to find two or three z scores of confidence interval:

> Q15. Can you find the z score of 95% confidence interval? Can you also find the z score of 90% and/or 99% confidence interval?

How to find t score of 90%, 95% and 99% confidence interval was not tested in this study because there are many t distributions with different sample sizes, and because t distribution was not emphasised in the participants' course.

9.3 Factors of difficulties

In the analysis of how participants looked for the z score of 95% confidence interval, I first found two main factors of their difficulties, including 'from $(1-\alpha)\%$ confidence interval to $Z_{\frac{\alpha}{2}}$', and 'memorised and recalled z score'. Then, when some participants attempted to find the z score of 90% or 99% confidence interval, I found another difficulty that arose when the probability was not inside the tables of normal distribution. In the following, I will discuss how three factors caused the participants' difficulties.

9.3.1 From $(1-\alpha)\%$ confidence interval to $Z_{\frac{\alpha}{2}}$

In Q15 of the first interview, the ten participants were asked to find the z score of 95% confidence interval by using the tables of normal distribution. I suggested that they not simply recall the memorised z score. This question was not displayed in a mathematical equation such as Q1.1-Q1.6 or Q3.1-Q3.2, so the probability for searching inside the Z table had to be developed by the participants themselves according to the given 95%. Earlier than Q1.a-Q1.d in the second interview, Q15 was the first time that the participants had to find a z score according to the given probability. As explained by Wonnacott and Wonnacott (1990) above, a probability of 0.025 rather than 0.05 should be used for searching the z score of 95% confidence interval. If the area 0.025 in the right-tail was obtained and searched for inside Table B, the question would be put in such a format:

Find $z_{0.025}$ if $P(Z > z_{0.025}) = 0.025$

However, only S1, S5, S9 and S10 attempted to get half of 0.05. Three of them correctly found 1.96 as the answer, but S1 mistakenly obtained 0.0025 rather than 0.025. The correct solving process from the narrative question to the z score is displayed in Figure 9.2.

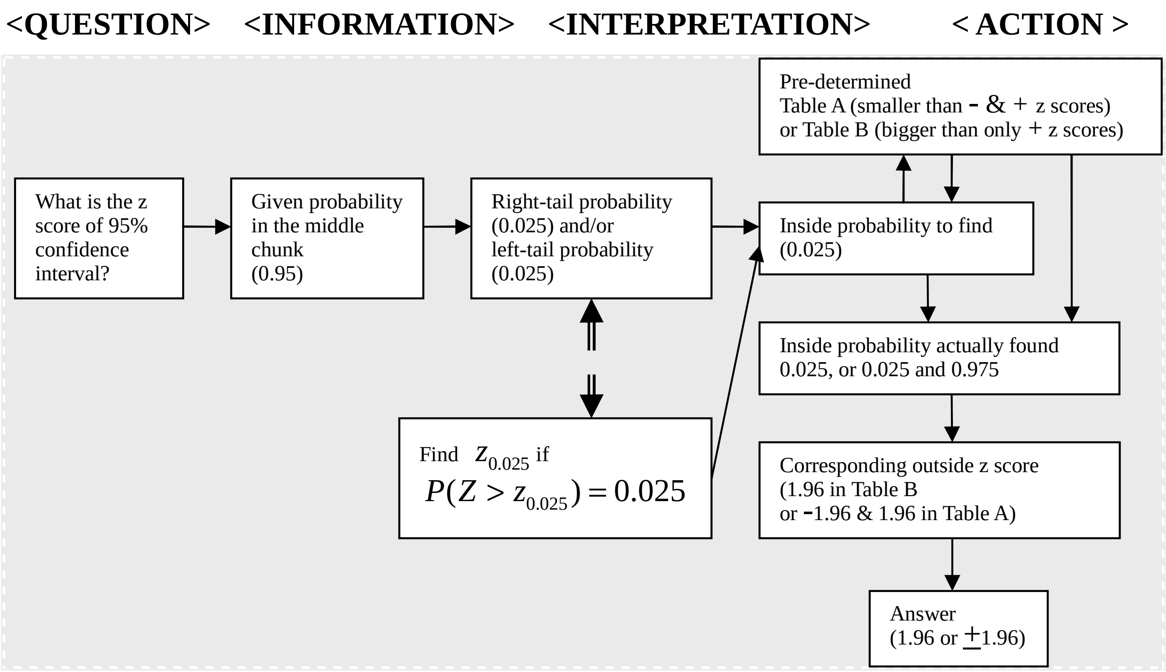


Figure 9.2. Solving processes of looking for z scores of 95% confidence interval.

S3 only obtained 0.05 for searching inside Table B. She explained that ‘‘there is not 0.95 in Table B, so I used 1 – 0.95’’ (S3, 29th Apr 09). Thus, she found the z score 1.645 as the z score of 95% confidence interval. She was confident in her answer because she was familiar with this number. However, she remembered 1.645 because it was the z score of 90% confidence interval. She solved this question as shown in Figure 9.3.

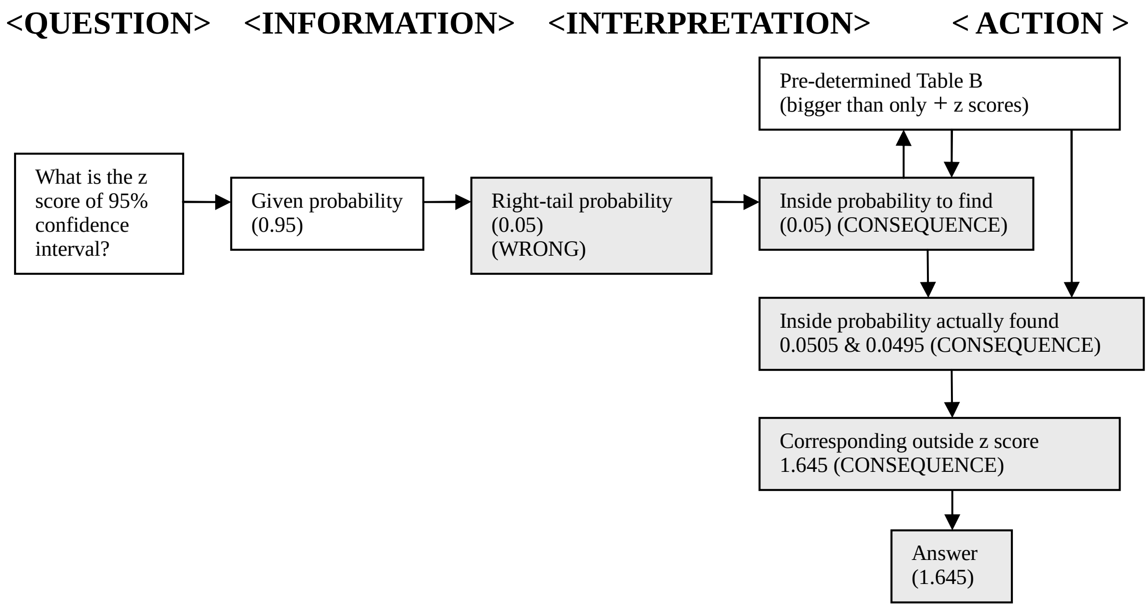


Figure 9.3. S3's solving processes of looking for z scores of 95% confidence interval.

S4 also tried to find 0.05 in Table B after she heard S3's explanation, but she found the z score 0.0 on the left-hand side and .05 on the top of Table B, rather than the inside probability. Such confusion of what is inside/outside the tables of normal distribution happened again when she solved Q1.a in the second interview. She believed that the answer should be in the column below .05, but she did not have a specific value.

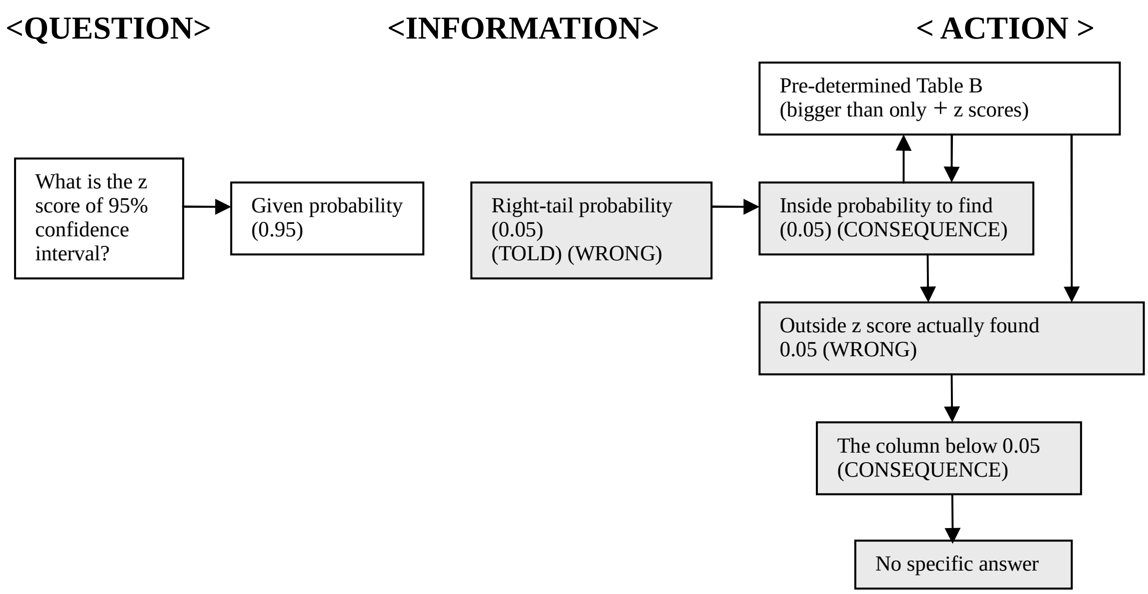


Figure 9.4. S4's solving processes of looking for z scores of 95% confidence interval.

I asked S6 to draw a graph of 95% confidence interval, and he drew a graph with confidence interval in the middle, obtained 0.025 for each tail and found 0.025 inside Table B and its corresponding z score outside Table B as the answer.

S2 and S8 kept trying to find 0.95, but they could not find 0.95 inside Table B, so they started using Table A. However, there was still no 0.95 inside Table A, and they did not consider .9495 and .9505. Therefore, they gave up trying to answer this question.

9.3.2 Memorised and recalled z scores

Although it was suggested to the participants that they not recall the three z scores when they solved Q15, some participants still relied on their memory. S1 explained that ‘‘it is easy to remember, so I did not think of how it was found’’ (S1, 29th Apr 09). However, the memorised z scores were not always correct (see the dialogues below for instance).

> I: … Can you find the z score corresponding to 95% confidence interval by yourself? Please use the table.
> …
> S7 could not find the z score in the table: I can memorise them, 90% is 1.575.
> I: Stop, stop.
> S8: 1.96.
> S7: 2.5.
> S6: 2.44, I also memorise them.
> (S6, S7, and S8, 4th May, 09)

S7 revised the recalled value 1.96 for 95% confidence interval, but she tried to find it inside Table B. Although the participants were told to find the z score, S7 regarded the probabilities .0197 and .1977 inside Table B as the closest values to 1.96. She appeared not to know what the recalled value 1.96 represented. Her mistake can be clearly seen in her solving steps (see Figure 9.5.).

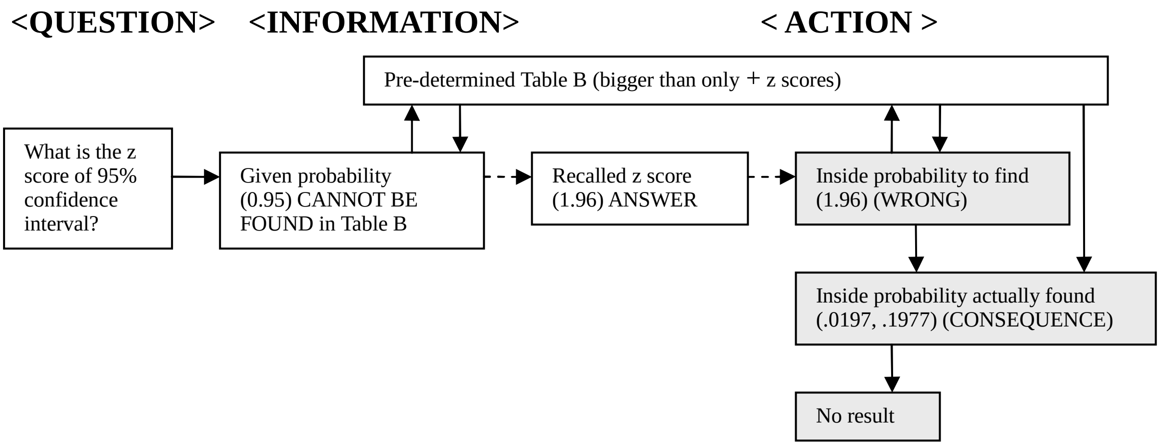


Figure 9.5. S7's solving processes of looking for z scores of 95% confidence interval.

** The dotted lines mean that S7 recalled the z score of 95% confidence interval as her answer and tried to find it inside Table B.*

After a discussion with other participants, S7 found the memorised z score 1.96 outside Table B, but she did not understand why the corresponding probability was not 0.05 but 0.025. Such a strategy of memorising without comprehension is a "superfluous didactic action" (Legutko, 2008, p.150).

Although S7 realised that the searched probability 0.05 was between 0.0505 and 0.0495 and that the corresponding z score of 90% confidence interval was 1.645 after a discussion with other participants, she still recalled 1.575 when looking for the z score of 90% confidence interval and could not find it inside Table B. The recalled z score 1.575 seemed to be a mix of the memorised 1.645 and 2.575. Her solving steps of looking for z scores of 90% confidence interval (see Figure 9.6) were similar to that of 95% confidence interval.

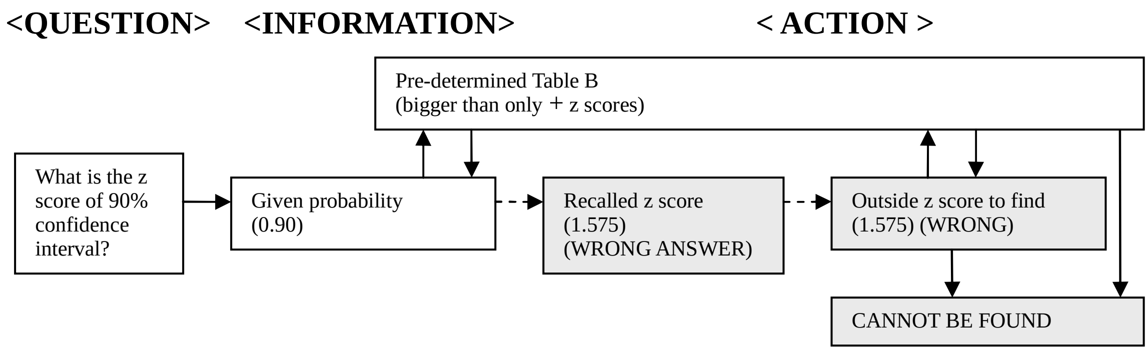


Figure 9.6. S7's solving processes of looking for z scores of 90% confidence interval.

** The dotted lines mean that S7 recalled the z score of 90% confidence interval as her answer and tried to find it inside Table B.*

In S7's case, she always recalled the z score directly before looking for probability in the table of normal distribution. S7's mistakenly memorised z score prevented her from looking at the correct values in the table. In other words, when she directly looked for the memorised z score in Table B, the probability that she should have looked for was not considered.

9.3.3 When the probability was not inside the tables of normal distribution

The z scores of 90% and/or 99% confidence interval were tested after the participants tried looking for the z score of 95% confidence interval. Five participants (S1, S5, S9 and S10, and S6 who found the answer after being suggested to draw the graph of 95% confidence interval) found 1.96 for 95% confidence interval by themselves. S6 and S10 also obtained 1.645 for 90% confidence interval and 2.575 for 99% confidence interval correctly, but S1, S5 and S9 had new problems. Their difficulty might have been caused by the probability which could not be found directly inside the tables of normal distribution.

S1 found 0.9949 and 0.9951 inside Table A when he looked for the z score of 99%

confidence interval, but he did not proceed or, perhaps, know how to proceed. S9's problem happened in interpreting the table. He found the probabilities 0.0051 and 0.0049 inside Table B and their corresponding z scores 2.57 and 2.58 when looking for the z score of 99% confidence interval, but he mis-interpreted the z score as 2.578 rather than 2.575 (see Figure 9.7.).

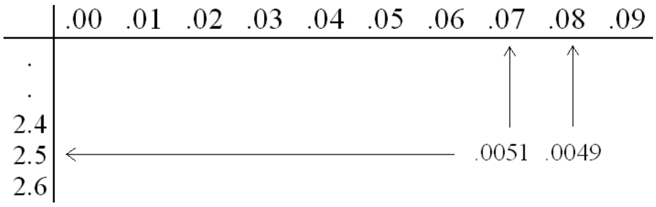


Figure 9.7 S9s' interpretation of z score 2.578

S5 obtained 0.025 (half of 5%) and found it inside Table B, as well as finding its corresponding z score 1.96 for 95% confidence interval. However, he did not attempt to find 0.05 (a half of 10%) inside Table B when looking for the z score for 90% confidence interval. He explained that 90% was close to 95%, so the z score of 90% confidence interval should be close to 1.96. Therefore, he believed that the z score of 90% confidence interval was the intersection of the row of 1.9 and the column of .05 (see Figure 9.8). His solving steps are explained in Figure 9.9.

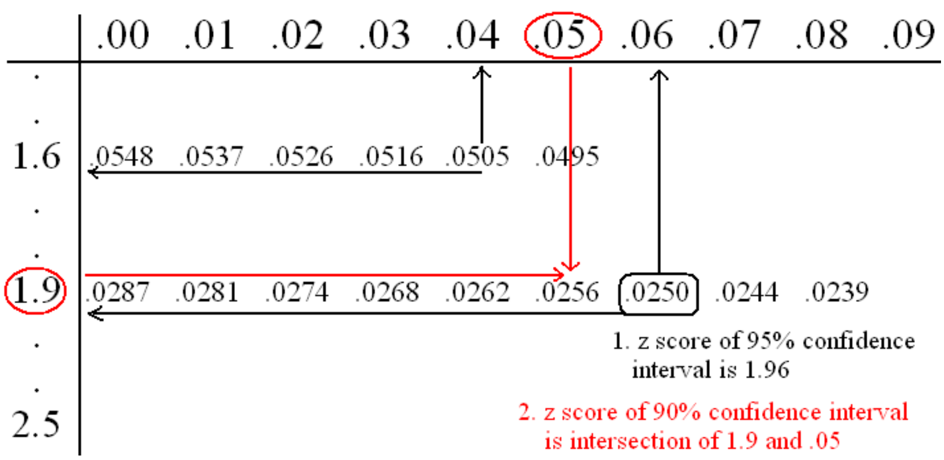


Figure 9.8. Comparison of S5's strategies to find z score of 95% confidence interval and

of 90% confidence interval.

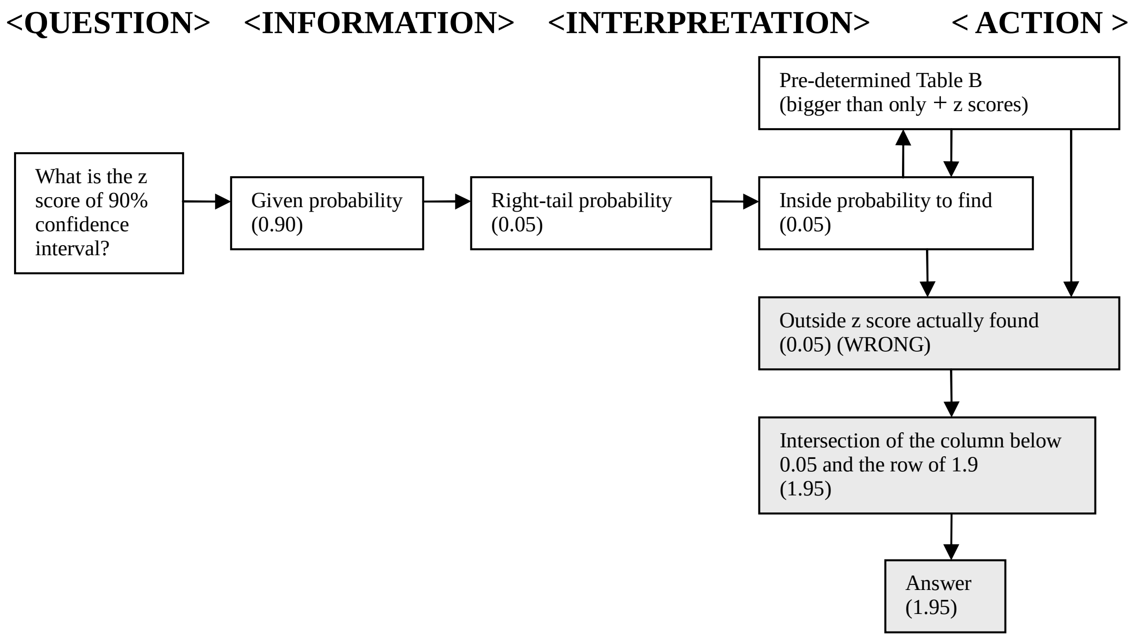


Figure 9.9. S5's solving processes of looking for z scores of 90% confidence interval.

Later S5 was reminded to find 0.05 inside Table B, but he could not distinguish 0.0548, 0.0537, 0.0526, 0.0516 and 0.0505 and said "they are all 0.05" (S5, 1st May 09). In the end, he thought that 0.0505 was the closest to 0.05 without considering 0.4995 since it was not in the form of 0.05xx. Therefore, he answered 1.64 as the z score of 90% confidence interval (see Figure 9.10.).

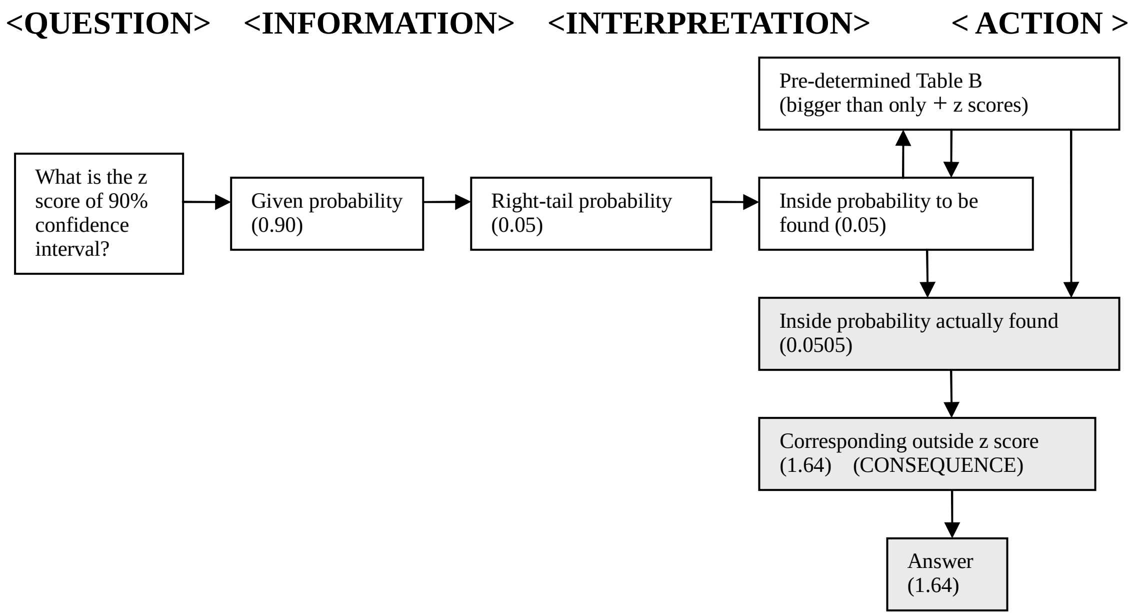


Figure 9.10. S5's solving processes of looking for z scores of 90% confidence interval.

After being prompted to find 0.05 for 90% confidence interval by other participants, S8 mis-interpreted the values inside Table B and thought that the answer was between 0.0051 and 0.0049 (as shown in Figure 9.11.).

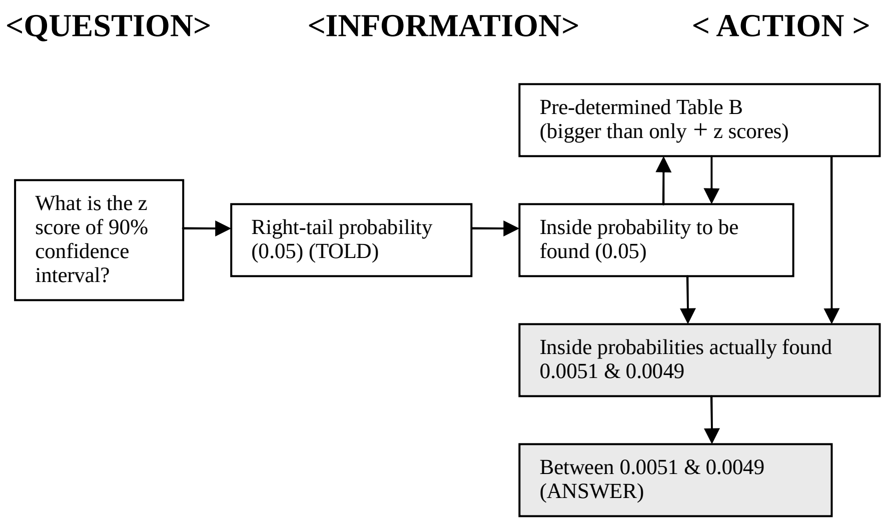


Figure 9.11. S8's solving processes of looking for z scores of 90% confidence interval.

As shown above, we can see that S1 S5, S8 and S9 had difficulties when the exact value could not be directly found inside the tables of normal distribution.

Three factors relating to the participants' difficulties in finding the z score of 90%, 95% and 99% confidence interval are discussed in this section. In the next section, I will also examine whether the participants memorised the z scores correctly.

9.4 How the memorising approach works

S6 and S7 actively recalled the memorised z scores when they had difficulty finding them in the tables of normal distribution. Other participants were also asked to recall the three z scores after they used the tables of normal distribution to find them. S2 and S4 did not answer; S9 did not answer because he believed that he was able to find the three z scores in Table B by himself. Whether and what others recalled are listed in Table 9.1.

	Recalled z score of 90% confidence interval	Recalled z score of 95% confidence interval	Recalled z score of 99% confidence interval
Correct	**1.645**	**1.96**	**2.575**
S1	Not asked	*0.96*	*2.757*
S3	1.645	1.96	2.575
S5	*1.945*	1.96	2.575
S6	Not asked	*2.44*… 1.96	Not asked
S7	*1.575*… 1.645	1.96	2.575
S8	Not asked	1.96	Not asked
S10	1.645	1.96	2.575

Table 9.1. Participants' recalled z score of 90%, 95% and 99% confidence intervals.

** 'Not asked' means that the participant was not asked to answer this question; '…' means the participant changed his/her answer; z scores in italics means the wrong answer.*

Table 9.1 shows that nearly one-third (5/16) of the participants' first recalled z scores were incorrect. Four of the wrong answers 1.945, 1.575, 0.96, and 2.757 looked similar to the correct answers, and the other wrong answer 2.44 seemed to be the wrong memorised value of 2.33 of 98% confidence interval (which was not used in the participants' statistics course). The students' memory was not as reliable as they had expected.

9.5 Standardisation (Normalisation)

Wild (2006) explains the meaning of standardised histograms of continuous measurement data. In such a format, proportions are represented by areas. This format involves probability density and its curve. Through standardisation, two or more data sets can be compared. The notions and the features of distributions draw on the behaviours exhibited in graphs of data, and statistical knowledge should be related conceptually to the pictures (2006, p.21). However, some participants attempted to find the z score according to the given values 95%, 90% and 99% without considering the meaning of probability and z score. The step of standardisation appeared to be just a formula for some participants. Therefore, I wondered how they understood the idea of distribution and its relationship to mean and standard deviation (or the variance), so I designed Q3 in the second interview to ask the participants to draw normal curves according to different means and standard deviations. The results indicate that some participants did not decide the centre of the normal curve according to mean, nor decide the spread according to the standard deviation (or variance). I will explain this in Section 11.4.

9.6 Participants' interpretation of confidence interval

In addition to the frequently used z score of 90%, 95% and 99% confidence interval, the meaning of confidence interval was also examined in the interview. The meaning of confidence interval was interpreted by the participants in different ways. S1 remembered the use of normalisation taught by the teacher, or that the confidence interval could be used to compare products and there would be a difference if the confidence interval (of difference of two means) did not include zero. His argument was an application of the rule that “a confidence interval may be regarded as just the set of acceptable hypotheses” (Wonnacott and Wonnacott, 1990, p.289). Once a confidence interval has been calculated, it can be used immediately to test any hypothesis.

However, S1 and S2 believed that 95% confidence interval was more precise than 90% confidence interval, and S2 explained that it was because 95% was closer to 100% than 90%. S3 and S5 did not explain the meaning of confidence interval, and they argued that 95% confidence interval was wider than 90% interval. S10 believed that the ‘‘value of error’’ or ‘‘range of floating’’ (S10, 4th May 09) of 95% confidence interval was smaller than that of 90%. The terms ‘value of error’ and ‘range of floating’ used by S10 were normally called “*error level*” (Wonnacott and Wonnacott, 1990, p.289). When a hypothesis is tested at a 95% confidence level, it is at an error level of 5% or with a 5% chance of error. I will discuss the application of confidence interval in hypothesis testing in Chapter 10.

S6 and S10 interpreted confidence interval as a range of estimation. They believed that the estimation would be wrong if it was outside the range, but they did not explain what it meant. S7 and S8 explained that confidence interval is ‘‘the range that we want, and it cannot be more or less’’ (S8, 4th May 09). S8 added that the range was where it may happen. These participants seemed to believe that confidence interval had to be the correct estimation without errors, and their interpretation of confidence interval was a lack of the idea of sampling.

9.7 Summary

Chapter 7 and Chapter 8 showed that the participants had several kinds of difficulty when using the tables of normal distribution and the table of t distribution to find a value according to the information given in symbolic format. This chapter investigated how the students found the three frequently used z scores of confidence interval. Some participants mistakenly memorised the z scores of 90%, 95% and 99% confidence interval without realising the relationship between the z scores and the probabilities, and they had problems finding the z scores of 90%, 95% and 99% confidence interval in the tables of normal distribution even if they recalled the memorised z scores. Such difficulties were not noticed in previous research, and they are deserving of more attention. In the next chapter, I will examine how the participants carried out the process of hypothesis testing, and my focus will be on how they found the critical value(s) and critical region(s).

CHAPTER 10

DATA ANALYSIS IV

CRITICAL VALUE AND CRITICAL REGION OF HYPOTHESIS TESTING

At the end of the second interview, the participants were given two questions of hypothesis testing of mean. However, some of them did not prepare for this part well, and five participants (S1-S5) agreed to attend a third interview after the final examination. The other five participants (S6-S10) could not attend for personal reasons. The two questions of hypothesis testing were tested again in the third interview without prior notice. However, hypothesis testing of population standard deviation σ will not be addressed in this study since the content was excluded by the teacher because of its difficulty. Therefore, the hypothesis testing discussed in this chapter will only focus on mean.

In this chapter, I will first introduce the three kinds of hypothesis and the approaches of hypothesis testing used by the teacher and the participants. I will then discuss the participants' solutions in the two exercise questions of hypothesis testing, such as their difficulty with making hypothesis, confusion of z score of confidence interval and z score of critical region, and inattention to the existence of negative critical region.

10.1 Making hypothesis: Two-tailed, left-tailed or right-tailed

In Wonnacott and Wonnacott's (1990) textbook, statistical hypothesis is defined as "a *claim about a population* that can be put to a test by drawing a random sample" (1990, p.288). There are three kinds of hypothesis, including two-tailed, left-tailed and right-tailed. Both left-tailed test and right-tailed test belong to one-sided test. Normally an individual needs to decide whether a two-tailed, left-tailed or right-tailed test should be undertaken according to the conditions.

> The one-sided test is appropriate when there is a one-sided claim to be made, such as "more than", "less than", "better than", "worse than", "at least", and so on… there are occasions when it is more appropriate to use a two-sided test. They often may be recognized by symmetrical claims such as "different from", "changed for better or worse", "unequal", and so on (Wonnacott and Wonnacott, 1990, p.314).

10.2 Approaches of hypothesis testing

In most statistics courses, one or some of the four approaches for hypothesis testing is/are taught: critical-value, test statistic, confidence interval, and p-value. The test statistic method was used by the teacher and the participants in this study. The p-value approach was also taught but rarely used by the teacher, and it was not used by the participants in the interviews. Here I will only briefly introduce the test statistic approach. A more detailed description and comparison of the four methods can be found in Appendix 10.1.

In the test statistic approach, no matter if the hypothesis is left-tailed, right-tailed or two-tailed, the test statistic of mean remains $Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$. The null hypothesis H_0 will be rejected only if the test statistic falls in the rejection region. The rejection region is determined by the type of hypothesis. More specifically, the rejection region in left-tailed testing is $\{Z < -Z_\alpha\}$, the rejection region in right-tailed testing is $\{Z > Z_\alpha\}$, and the rejection region in two-tailed testing is $\{Z < -Z_{\frac{\alpha}{2}} \text{ or } Z > Z_{\frac{\alpha}{2}}\}$. This approach was taught and used by the teacher, and it was also used by all participants to solve the two questions of hypothesis testing in the second interview and the two repeated questions in the third interview.

10.3 Discussion of the participants' solutions in the two questions of hypothesis testing

Below I will discuss the participants' solutions in the two similar narrative questions of hypothesis testing. The two questions were Q6 and Q7 in the second interview. They were tested again as Q9 and Q10 in the third interview.

> Q6. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the content is 500c.c. and write down your steps. α = 0.05

> Q7. It is said by a drink seller that the content of their drink is over 500c.c. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10c.c. Please test whether the drink seller is telling the truth, and write down your steps. α = 0.05

The only difference between the two questions is that Q6 was two-tailed (equal to), and that Q7 was right-tailed (bigger than). Five participants (S1-S5) attended the third interview and answered Q9 and Q10 (see Appendix 6.18.). There were a few problems found in their solutions, such the participants' difficulty with making hypothesis, confusion of z score of critical value and z score of confidence interval, and inattention to negative critical region in two-tailed testing and left-tailed testing. I will discuss each difficulty below.

10.3.1 The participants' difficulty with making hypothesis

The first difficulty was that some participants did not make the correct hypotheses according to the narrative questions, such as S3, S9 and S10 making left-tailed hypotheses and S6 making right-tailed hypotheses in Q6, which was a two-tailed situation. Furthermore, sometimes the coverage of null hypothesis and alternative hypothesis was not complementary. For example, S1 made hypotheses in Q6 as:

$$\begin{matrix} H_0 = 500 \\ H_1 < 490 \end{matrix}\text{; and}$$

S5 and S8 made hypotheses in Q6 as:

$$\begin{matrix} H_0 : \mu = 500 \\ H_1 : \mu < 500 \end{matrix}\text{; and}$$

In Q6 of the second interview, eight participants made hypotheses (S2 and S4 were the two exceptions). Three of them (S1, S5 and S8) made a null hypothesis and an alternative hypothesis, S6 made hypotheses for right-tailed testing, and the other three participants made hypotheses for left-tailed testing. Therefore, only S7 correctly made hypotheses for two-tailed testing in Q6.

On the other hand, five participants correctly made hypotheses for left-tailed testing in Q7. It seemed that it was easier for the participants to make hypotheses in left-tailed testing than to make hypotheses in two-tailed testing according to the narrative question. Their hypotheses in Q6 and Q7 are listed in Table 10.1.

	Q6 (the content is 500, average 498.2)	Q7 (the content is over 500, average 498.2)
correct	**Two-tailed,** $H_0: \mu = 500, H_1: \mu \neq 500$	**Left-tailed,** $H_0: \mu \geq 500, H_1: \mu < 500$
S1	$H_0 = 500, H_1 < 490$	$H_0 < 500, H_1 > 500$
S3	$H_0: \mu \geq 500, H_1: \mu < 500$	
S5	$H_0: \mu = 500, H_1: \mu < 500$	$H_0: \mu \geq 500, H_1: \mu < 500$
S6	$H_0: \mu \leq 500, H_1: \mu > 500$ right-tailed testing	$H_0: \mu \geq 500, H_1: \mu < 500$ left-tailed testing
S7	$H_0: \mu = 500, H_1: \mu < 500$ or $H_1: \mu \neq 500$ two-tailed testing	$H_0: \mu \geq 500, H_1: \mu < 500$
S8	$H_0: \mu = 500, H_1: \mu < 500$	$H_0: \mu \geq 500, H_1: \mu < 500$
S9	$H_0: \mu \geq 500, H_1: \mu < 500$	$H_0: \mu \geq 500, H_1: \mu < 500$
S10	$H_0: \mu > 500, H_1: \mu < 500$	

Table 10.1. S1-S10's hypotheses in Q6 and Q7 in the second interview.

**Hypotheses made by participants were parallel. They were listed in this way for convenience to compare.*

***S2 and S4 did not make hypotheses in Q9 or Q10.*

S1-S5 attended the third interview after the final examination. The two questions were tested again as Q9 and Q10 in the third interview without prior notice, and the hypotheses they made in the two questions are also listed in Table 10.2.

	Q6 (the content is 500, average 498.2)	Q7 (the content is over 500, average 498.2)
Correct	**Two-tailed,** $H_0: \mu = 500, H_1: \mu \neq 500$	**Left-tailed,** $H_0: \mu \geq 500, H_1: \mu < 500$
S1	$H_0: \mu = 500, H_1: \mu \neq 500$	$H_0: \mu > 500, H_1: \mu \leq 500$
S3	$H_0: \mu = 500, H_1: \mu < 500$	$H_0: \mu > 500, H_1: \mu < 500$
S4	$H_0: \mu \leq 100, H_1: \mu > 100$	$H_0: \mu \geq 100, H_1: \mu < 100$
S5	$H_0: \mu = 498.2, H_1: \mu \neq 498.2$	$H_0: \mu \geq 500, H_1: \mu < 500$

Table 10.2. S1-S5's hypotheses in Q9 and Q10 in the third interview.

**Hypotheses made by participants were parallel. They were listed in this way for convenience to compare.*

***S2 did not make hypotheses in Q9 or Q10.*

In the third interview, S2 still did not make hypotheses, and the other four participants made hypotheses for left-tailed testing in Q10, but S4 regarded the given μ as the sample size. However, most of them still had difficulties in making hypotheses for two-tailed testing according to the narrative question. This difficulty was consistent with the finding in the second interview.

After reviewing the participants' textbook, I found that their in-complementary hypotheses seemed to be influenced by the examples in their textbook, such as

e.g. 9-1.8: $H_0 : \mu = 800, H_1 : \mu > 800$;

e.g. 9-1.13: $H_0 : \mu = 100, H_1 : \mu > 100$; and

e.g. 9-5.5: $H_0 : \mu_1 - \mu_2 = 7, H_1 : \mu_1 - \mu_2 > 7$ (Kuo and Shi, 2006, pp.327, 335)

The teacher also used these hypotheses in examples in class, but he did not notice any problem with the hypotheses. Therefore, some participants made such hypotheses in the interviews, and this influenced their solutions. In fact, such hypotheses are also made in other textbooks, such as

$$\begin{aligned} &H_0 : \mu = 1200 \\ &H_A : \mu > 1200 \end{aligned}$$ (Wonnacott and Wonnacott, 1990, p.293)

Although Wonnacott and Wonnacott (1990) explain in this case that they "use just the boundary point $\mu = 1200$ " (1990, p.293) to make $H_0 : \mu = 1200$ rather than

$H_0: \mu \leq 1200$, and “[a]ny other μ below 1200 would be easier to distinguished from H_A since it is “further away”” (1990, p.293), such hypotheses may influence students’ learning. Such influence has been seen in this study. According to this finding, we can see that the impact of the textbook on students’ and teachers’ errors should be more carefully considered (Legutko, 2008).

10.3.2 The participants’ confusion of z score of critical region and z score of confidence interval

When the participants dealt with the two questions of hypothesis testing, some participants’ concept of α (significance level of hypothesis testing) appeared to be highly influenced by their concept of $1-\alpha$ (confidence coefficient or degree of confidence of confidence interval). Some participants could not distinguish them and regarded them as always the same, but critical region is not necessarily the same as confidence interval.

Indeed, there is an equivalence of the two-sided critical region approach and the two-sided confidence interval approach, and also an equivalence of the one-sided critical region approach and the one-sided confidence interval approach (see Appendix 10.1.). The difference between the two methods is that the critical region approach uses the assumed population mean μ_0 as its reference point, while the confidence interval uses the sample mean $\overline{X}$ as its reference point (Wonnacott and Wonnacott, 1990), but both methods involve z score. However, some participants who simply used the memorised positive z score of confidence interval when looking for the z score of critical value might have had a problem, because they did not consider whether the testing was two-tailed, left-tailed or right-tailed. The participants (and many other students) only learned two-sided confidence interval and did not consider the influence of type of testing on the z score. Such inattention to the influence of the condition on the z score was also found in Q15 in the first interview.

In the second interview, some participants did not answer Q6 and Q7, and some of them did not have the complete process of solution but rather developed their own strategy (see Appendix 10.2.). For example, S1 compared $\mu_0 \times \alpha$ and s (sample standard deviation), and S2 compared $\overline{X} - \mu_0$ and s. Only S5 and S6 had complete solving processes in Q6 and Q7 with the test statistic method. S7, S8 and S10 had the partial solution with the test statistic method. S9 also used the test statistic method, but he used t score rather than z score even if the sample size was 100. The participants' critical values in Q6 and Q7 are listed in Table 10.3.

Question / Critical value	Q6 ($\alpha = 0.05$)	Q7 ($\alpha = 0.05$)
Correct	**Two-tailed** $\pm z_{\frac{\alpha}{2}} = \pm z_{\frac{0.05}{2}} = \pm z_{0.025} = \pm 1.96$	**Left-tailed** $-z_{\alpha} = -z_{0.05} = -1.645$
S5	2.575 (by memory)	2.575 (by memory)
S6	$z_{0.025} = 1.96$	$-z_{0.025} = -1.96$
S7	-1.96 and 1.96 (by memory)	-1.96 (by memory)
S8	$z_{0.05} = 1.645$	
S9	$t_{0.05}(99) = 1.671 \sim 1.658$	$t_{0.05}(99) = 1.671 \sim 1.658$
S10	$z_{0.05} = 1.645$	

Table 10.3. Correct critical values and S1-S10's critical values in Q6 and Q7 in the second interview.

**Blank means that no critical value was used in their solution.*

***S1-S4 did not find or use critical value in Q6 or Q7.*

Some participants directly took the memorised z score of confidence interval (i.e. 1.645 for 90%, 1.96 for 95%, or 2.575 for 99%) to be the critical value in Q6 and Q7 without considering the type of testing. For instance, S5 used 2.575 as the critical value in both questions with $\alpha = 0.05$. He explained that the testing processes of Q6 and of Q7 “should be the same” even if he noticed that the condition in Q6 was “equal to 500” and the condition in Q7 was “bigger than 500” (S5, 5th June 09). S8 and S10 did not recognise Q6 as a two-tailed testing, so they used $z_{0.05} = 1.645$ rather than $z_{0.025}$ as the critical value. However, they did not find the critical value in Q7.

S7 believed that Q6 was a two-tailed testing, and she was the only one who correctly selected -1.96 and 1.96 as the critical values in Q6. However, she still used -1.96 as the critical value in Q7 even though she knew that Q7 was a left-tailed testing. S6 regarded Q6 as a right-tailed testing with $\alpha = 0.05$, but he used $z_{0.025} = 1.96$ as the critical value. Similarly, he regarded Q7 as a left-tailed testing with $\alpha = 0.05$, but he used $-z_{0.025} = -1.96$ as the critical value.

S1-S5’s critical values in Q9 and Q10 in the third interview are listed in Table 10.4. S3 did not answer Q6 and Q7, and her critical values in Q9 and Q10 seemed to be the recalled memorised z scores of confidence interval.

Question Critical value	Q9 (two-tailed)	Q10 (left-tailed)
Correct	$\pm z_{\frac{\alpha}{2}} = \pm z_{\frac{0.05}{2}} = \pm z_{0.025} = \pm 1.96$	$-z_{\alpha} = -z_{0.05} = -1.645$
S1	-1.96	~~-1.96~~, -1.645
S3	90 95 → 0.05 99 → 0.1 $z_{0.05} =$ ~~2.575~~, ~~1.96~~, -1.96	$z_{0.05} =$ ~~2.575~~, -2.575
S4	$z_{0.05} \rightarrow 1.645$	
S5	1.96	1.96

Table 10.4 Correct critical values and S1-S5’s critical values in Q9 and Q10 in the third interview

***S2 did not find or use critical value in Q9 or Q10.*

By comparing S3's processes to find the critical values in Q9 and Q10 with the correct z score of confidence interval that

90% $z_{0.05} = 1.645$

95% $z_{0.025} = 1.96$

99% $z_{0.005} = 2.575$,

we can find two problems in her steps to find the critical values. First, S3's symbolic representation of critical value did not match the situation of Q9 and Q10. In both questions, she took $z_{0.05}$ as the critical value according to $\alpha = 0.05$ without considering how the type of testing influenced the probability in the symbolic representation. Similar to her solution in Q15 in the first interview (finding z score of confidence intervals), S3 had difficulty getting the correct probability for searching in the tables of normal distribution.

Second, S3 obtained the critical value by recalling the memorised z score of confidence interval without using the tables of normal distribution, since she wrote in Q9 that

$\frac{\alpha}{2} = 0.1$ (which was the result of mis-division of $\frac{0.05}{2}$. The correct result was 0.025)

and

90

95 → 0.05

99 → 0.1 (the result of mis-subtraction of 1 – 0.99)

Therefore, S3 initially got $z_{0.05} = 2.575$ (which was the z score of 99% confidence interval) for Q9. Later, she found the correct result of $\frac{\alpha}{2} = 0.025$ and revised the critical value to $z_{0.05} =$ ~~2.575~~ 1.96. However, she did not revise $z_{0.05}$ to $z_{0.025}$. Although S3 had the correct idea of how two-tailed testing and one-tailed testing influenced the probability of the tail, she might still have had difficulty representing the probability in symbolic form. Also, she preferred recalling the memorised z scores of confidence interval to finding the z score in the tables of normal distribution according to the probability (see the dialogues below).

> I: In Q10, do you accept or reject H_0?
> S3: To accept H_0. $H_0 : \mu > 500$, so it is one-tailed testing. $\alpha = 0.05$, so the critical value is 2.575…
> (S3, 25th Jun, 09).

For some participants, looking for the critical value was difficult since they could not correctly make two-tailed, left-tailed or right-tailed hypotheses and get the correct probability of the tail(s). Moreover, many participants could not get the positive and/or negative critical value(s) $z_{\frac{\alpha}{2}}$ or z_{α} according to the narrative questions. I will explain this in Section 10.3.3.

10.3.3 Disregarded negative critical region in two-tailed testing

Some participants did not consider that the critical value could be negative even if they made an alternative hypothesis for left-tailed testing. For instance, S5, S8, S9 and S10 made the alternative hypothesis $H_1 : \mu < 500$ in Q6, but their critical value in Q6 was still positive.

Comparing S1-S5's performance in Q9 and Q10 in the third interview and their performance in Q6 and Q7 in the second interview, some of them made progress and became more capable of solving the two questions (see the table of Appendix 6.19.). They also used the test statistic approach, so they needed to find the critical values. S2 found Z_0, but he did not find critical values. Instead, he compared Z_0 to the standard deviation. S2 misinterpreted the symbol of $\overline{X}$ and μ, so he obtained a positive test statistic:

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{500 - 498.2}{\frac{10}{\sqrt{100}}} \quad \text{(the correct } \mu \text{ was 500 and } \overline{X} \text{ was 498.2)}$$

Initially, S3 also reversely interpreted the symbol of $\overline{X}$ and μ, obtained a positive test statistic Z_0 = 1.8 in Q9 and compared it to the critical value $z_{0.05} =$ 2.575. When she solved Q10, she recognised Q9 as two-tailed testing and Q10 as one-tailed testing, so she revised the critical value $z_{0.05} = \cancel{2.575}$ 1.96 in Q9 and gave the critical value $z_{0.05} = 2.575$ in Q10. Soon she found her mis-interpretation of $\overline{X}$ and μ, after which she got the correct Z_0 = -1.8. She then compared the negative value of test statistic Z_0 = -1.8 to negative critical values -1.96 in Q9 and -2.575 in Q10. She explained that "[in Q10]... the critical value is 2.575, but it is negative because Z_0 is negative" (S3, 25th Jun 09).

S4 regarded the sign μ as sample size, so she had an incorrect Z_0. She did not consider Q9 as a two-tailed testing and found the critical value $z_{0.05} = 1.645$.

S5 correctly recognised Q9 as two-tailed testing and Q10 as left-tailed testing, and he got the negative Z_0 (-1.8). However, he did not consider that there were two critical values of two-tailed testing and still compared the negative Z_0 to the positive critical value 1.96. He explained that the critical value 1.96 came from the significant level 0.05. In S5's case, recognising the question as two-tailed testing or left-tailed testing did not influence his choice of critical value and did not lead him to consider the existence of negative critical values.

Therefore, only S1 was able to get the correct critical values and correctly solve Q9 of two-tailed testing and Q10 of left-tailed testing in the third interview.

10.4 Summary

According to the solutions of the participants who decided critical values (i.e. S5-S10 in Q6 and Q7 in the second interview and S1 and S3-S5 in Q9 and Q10 in the third interview), we can find a few difficulties in their answers. First, most participants could not recognise Q6 or Q9 as a two-tailed testing according to the narrative question with 'the content is 500c.c.', so they treated it as a one-tailed testing.

Second, most participants' critical values in both questions were the same. They simply took the critical value found in Q6 or Q9 as the critical value of Q7 or Q10. Even if some participants distinguished Q6 or Q9 as a two-tailed testing and Q7 or Q10 as a left-tailed testing, they still used the same values such as '-1.96 and 1.96', or '-1.96' as the critical values in both questions. In other words, most participants decided the critical value(s) without considering the influence of the type of testing.

Third, some participants' critical value in test statistic method remained positive even if they made a left-tailed alternative hypothesis or obtained the negative test statistic Z_0 by calculation. The participants' inattention to negative numbers was observed in Q3.2 and Q1.a.

I have analysed and discussed the participants' solutions to the exercise questions in Chapters 7, 8, 9 and 10. In the next chapter, I will examine the participants' use of and attitude towards visual representation (i.e. graphs) in solving these questions. I will also discuss how they drew graphs of normal curves, which reveals their interpretation of the relationship of mean, standard deviation and the normal curve.

CHAPTER 11

DATA ANALYSIS V

VISUALISATION

11.1 Introduction

In Chapter 7 to Chapter 10, I presented an analysis of the participants' solutions to the exercise questions concerning the use of the tables of normal distribution and t distribution, how they looked for z score of confidence interval and critical value(s), and the critical region(s) of hypothesis testing in the three clinical interviews. The focus of this chapter will be on the use and influence of visualisation (in specific, graphs of distribution) in the participants' solutions to these questions.

In the beginning of this chapter, I will briefly introduce the participants' interpretation of and attitude towards visualisation. Then I will discuss the specific application of visualisation (i.e. the graphs of normal distribution and t distribution) in this study and the benefits of using visual aids in the part of statistical inference. I will also explain the reasons for and consequences of the participants' reluctance to use graphs, and I will compare the solutions of participants who used and did not use graphs. The graph users performed better in transferring steps than the graph non-users. At the end of this chapter, I will examine how the participants drew normal curves according to symbolic representation. The result reveals their thoughts on how the mean and standard deviation influence the construction of the normal curve.

11.2 Visualisation in this study

11.2.1 Participants' interpretation of and attitude towards visualisation

In the first interview, the participants described their interpretation, attitude and actual use of visualisation in their statistics learning. Most participants interpreted the term visualisation as graphs or drawing graphs. S3 had a broader definition of visualisation: "what you see through your eyes, including graphs, formulas or theorem" (S3, 29th Apr 09). S5 included a table in his definition of visualisation, but he paid more attention to how the teacher solved the questions with formulas. He admitted that "we imitate methods used in the previous question without really understanding" (S5, 1st May 09).

Duval (1999) suggests that both symbolic representation and visualisation are at the core of mathematics understanding. Visualisation can be a "tool for meaning-making in mathematics" (Breen, 1997, p.97). For instance, S6 shared his experience learning and using visual forms as follows:

> I spent much time to think of what the graph is about… when I see some graphs, I know that they are very difficult so I hate them, but I am still used to drawing graphs in the beginning when solving some questions (S6, 4th May 09).

S7 and S8 felt the necessity of graphs to explain ideas of intersection and union of sets, but they did not like nor feel the necessity of the graphs of normal distribution and t distribution. S7 believed that

> you can solve questions without drawing graphs… drawing graphs is troublesome and unnecessary" (S7, 4th May 09).

S2, S4 and S8 felt that the graphs of normal distribution and t distribution in statistical inference were too difficult to learn and use. S9 and S10 relied totally on formulas to solve questions and did not attempt to use any graph. Therefore, six participants did not actively use any graph in the entire study.

The participants were not requested to use any particular method or aid except for the tables of distribution to solve most exercise questions. However, two participants drew the graph of curve with shadow area(s) in left-tail, right-tail or both tails, or used the attached graphs above the tables of normal distribution or t distribution in their textbook to help them solve the exercise questions. Two more participants also drew graphs in the second and third interviews.

11.2.2 Terminology

People who use visual representations are called “visualisers” (Lowrie, 2001, p.354; Presmeg, 2006, p.206), and people who prefer verbal methods are called verbalisers or non-visualisers. In this study, most exercise questions were in symbolic form, and other questions were in narrative form. The participants used or manipulated symbolic forms (verbal approach) in all exercise questions except for Q3 in the second interview which asked them to draw normal curves according to the given means and standard deviations. Some participants also used graphs (visual approach) occasionally, but none of them only used graphs without manipulating symbolic equations, and none of them used both approaches in all questions. For instance, S6 used graphs only when he felt the question difficult. That is, only a few participants used both approaches in some questions.

The tables of normal distribution (Appendix 2.1. or/and Appendix 2.2.) and the table of t distribution (Appendix 2.3.) were needed for the participants to find particular probability value, z score or t score according to the information, perhaps with some subtractions.

The tables used in the study link to the structure of graphs with the function to display data or organize information as a “transition tool” (Friel *et al.*, 2001, p.128) to create graphical representations, and both tables and graphs are included in visualisation. Therefore, in order to prevent readers from interpreting visualisers as people who use tables, I will name the participants who attempted to take advantage of graphs in the exercise questions “graph users” instead of visualisers in this study.

Graph users do not necessarily draw the whole graph by themselves. For example, S6 simply marked -0.03 on the left-end of the x-axis in the graph above Table A (see Figure 11.1.) but he did not make a shadow area when he solved Q1.1.

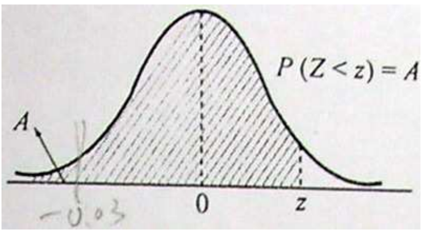


Figure 11.1. S6’s use of the graph above Table A for Q1.1.

He also marked -3.1 on the left end of the x-axis in the graph above Table B and made a shadow area in the right-side of the border when he solved Q1.3 (see Figure 11.2.). He drew graphs by himself in Q1.5 and Q1.6.

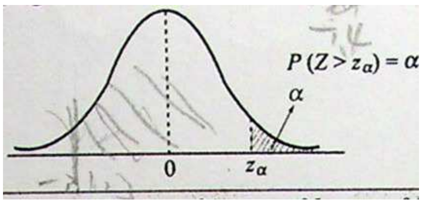


Figure 11.2. S6’s use of the graph above Table B for Q1.3.

S5 also said that he took advantage of the graphs attached in the textbook because he had difficulty drawing graphs for new questions by himself.

11.2.3 Elements of graph and table of distributions

According to Duval (1999) and Arcavi (2003), realising visualisation as a whole story requires consideration of some characteristic features, such as the type of curve, direction of shadow area, size of shadow area and location of the z or t score in the case of graphs of normal distribution and t distribution. Therefore, the graphs of normal distribution should include the three main elements in the equations: the direction of the shadow (relation sign), the value of the border (z score) and the shadow area (probability). In most situations, the direction and one of the other two elements were implied by the exercise questions, such as $P(Z > 0.67) = b$ (right-tail and z score were given) and $P(Z < c) = 0.67$ (left-tail and probability were given). The given information can be shown in the graph and lead to the unknown value.

However, the graphs of t distribution do not include the information of degree of freedom, which is given in the symbolic representation, such as $t_{0.005}(9)$ or $t_{0.99}(23)$. This is one of the limitations of the graph of t distribution. In addition, the right-tail characteristic of the graph of t distribution was not presented in the symbolic form, so the users needed prerequisite knowledge to correctly interpret the symbolic representation.

The total area 1 below the normal curve or the t curve was missing in the graphs of normal distribution and t distribution. Lack of indication of total area 1 in the graphs might be the reason of why some graph users could not take advantage of the idea of 'complementary' of the shadow area and the other area. Such a problem was first explored in Q3.2, and was also found in some other questions.

11.2.4 Benefits to using graphs in the study

S1 and S3 believed that graphs enabled them to compare 'bigger than' and 'smaller than' because "pure numbers are chaotic… with graphs, I can better know which place the number should be at" (S1, 29th Apr 09).

S3 took advantage of graphs to enhance her memory because she believed that she could remember graphs better than formulas. She felt it more difficult to memorise the formulas (i.e. transferring principles) without errors and easier to remember graphs, so she wanted to indirectly "remember the formula by remembering the graphs" (S3, 29th Apr 09).

Graphical representation can more explicitly preserve information about geometric relations among the components of the questions than sentential representation (Larkin and Simon, 1987). The graphs of the normal curve not only helped explain the transferring principles but also offered the elements (i.e. corresponding z score and probability) of the tables of normal distribution. As suggested by Guzmán (2002), the image is "frequently the matrix from which concepts and methods arise" (p.10).

11.2.5 Reluctance to use graphs

The teacher used graphs to explain the ideas of transferring principles of using the tables of normal distribution, and although he occasionally drew graphs to explain how he solved the questions in class, some participants simply used the table of normal distribution to find the given z score without transferring, and some others recalled the memorised principles and used the tables of normal distribution to find the obtained z score after transferring without using any graph. Mundy (1984) also finds that

> [some] students did not have a visual understanding that integrals of positively

valued functions can be thought of in terms of an area under a curve (Mundy, 1984, cited in Eisenberg and Dreyfus, 1991, p.28).

Eisenberg and Dreyfus (1991) explain that students prefer algorithmic to visual thinking because "thinking visually makes higher cognitive demands than thinking algorithmically" (1991, p.25):

> The analytic argument is short, neat, assumes little and gives the result without any of its wider implications. It is easy for the students to learn, easy to apply to exercises… The visual argument… builds on prerequisite visual knowledge… It shows additional, related, but not essential, information… it is thus more difficult to understand (Eisenberg and Dreyfus, 1991, p.30).

Eisenberg and Dreyfus (1991) explain that students had difficulties because "they have not learned to exploit the visual representations associated with the concepts (1991, p.25). The visual interpretations, even known by the students, are not in the core of knowledge and not exploited in the students' thinking processes. Therefore, memorising transferring principles and manipulating the symbols in the formula seemed to be easier than drawing graphs for some students who might not possess adequate knowledge of graphs. Thus, eight participants did not use any graph in the first interview, and six participants did not actively use any graph in the second and third interviews.

11.3 Graph users' and graph non-users' performances in the exercise questions

The exercise questions in the clinical interviews can be categorised into five groups, namely using the tables of normal distribution to find a value (probability) in Q1.1-Q1.6, using the table of t distribution to find a value (t score) in Q3.1-Q3.2, using the tables of normal distribution to find z scores of confidence interval in Q15, using the tables of normal distribution to find a value (probability or z score) in Q1.a-Q1.d, and in a part of the overall solving procedures, using the tables of normal distribution to find critical value(s) and critical region(s) of hypothesis testing in Q6, Q7, Q9 and Q10.

Only two participants, S3 and S6, used graphs in all kinds of questions (but not in each question), and two others, S5 and S7, used graphs in the later kinds of questions. Except for Q15, the type of unknown value was implied by the symbolic representation of each exercise question, and some participants did not interpret the symbolic form correctly, as discussed in previous chapters. This might influence the correctness of the graphs. In this section, I will compare the graph users' and graph non-users' performance in each type of question, examine the influence of the graphs and analyse what problems graph users might have.

11.3.1 Graph users, and graph non-users' performances in Q1.1-Q1.6

S3 and S6 drew graphs in Q1.5 and Q1.6. S6 also drew graphs in Q1.1 and Q1.3. Despite some careless mistakes in the solving processes, S3 and S6 both completed the thirteen transferring procedures for Table B users in Q1.1-Q1.6 without any difficulty (see Appendix 7.1). Their solving processes in questions in basic format Q1.1-Q1.4 were very similar to the standard solving steps (see Figure 5.2).

When S3 and S6 solved Q1.5 and Q1.6, they first drew graphs (see Table 11.1.) and transferred the question of finding the probability of Z between two z scores to the difference of two probabilities of Z being bigger than a z score (see Appendix 7.1. and Figure 5.3.), and they then dealt with the two basic forms. Graphical representation can better preserve information about geometric relations among the components of the questions (Larkin and Simon, 1987). For instance, S3 pointed at her self-drawn graph in

Q1.6 and clearly explained that

> the middle area can be obtained from the difference of subtracting this area from the whole area and subtracting that area from the whole area (i.e. [1 – P(Z > 2.41)] – [1 – P(Z > 0.03)]) (S3, 29th Apr 09).

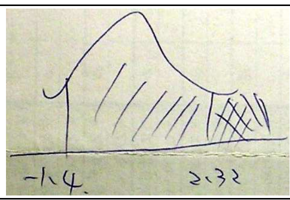
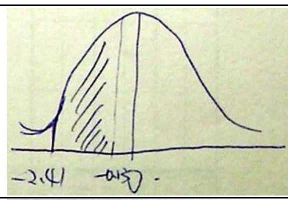
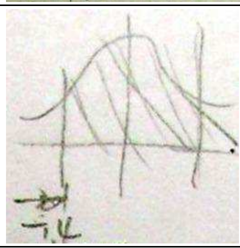
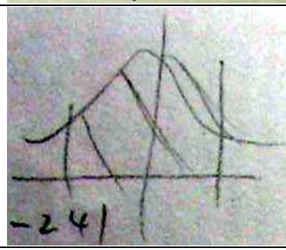
	Q1.5	Q1.6
	P (-1.4 < Z < 2.32)	P (-2.41 < Z < -0.03)
S3		
S6		

Table 11.1. S3's and S6's graphs in Q1.5 and Q1.6.

On the other hand, some graph non-users also transferred well, while some others did not. For example, S5's and S10's transferring steps in Q1.1-Q1.6 were all successful. S9 had problems in transferring 'smaller than' to 'bigger than'. S1 used Table B, which he was not familiar with, to solve the questions, and this perhaps caused his wrong transferring. Their solving steps in Q1.1-Q1.4 and Q1.5-Q1.6 were similar to those of S3 and S6, but the graph and its influences were excluded (see Figure 5.2. and Figure 5.3.).

Guzmán (2002) argues that what is helpful (such as graphs) for one person may be a hindrance for another, and graphical representations seemed to be unnecessary for some participants since they relied on symbolic manipulation to solve the questions correctly.

However, some participants were found unable to overcome some particular questions without involving graphs. For instance, S2 neither understood nor remembered the transferring principles of normal distribution, and he did not learn to use the graphs. His strategy was to follow the steps in the examples in his textbook, but he could not distinguish the types of questions. Taking Q1.5, $P(-1.4 < Z < 2.32) = ?$ as an example, S2 directly found two probability values 0.0808 and 0.0102 corresponding to 1.4 (the minus sign was neglected) and 2.32 in Table B respectively, and he then attempted to subtract 0.0808 from 0.0102 in order to "get the middle" (S2, 29th Apr 09). However, he did not know how to proceed when he was told that '0.0102 - 0.0808' was negative.

S7 performed well in transferring, but she did not use graphs because she felt drawing graphs to be "troublesome and unnecessary" (S7, 4th May 09). Completely relying on symbolic manipulation, S7 made a unique mistake in Q1.5 in which $P(-1.4 < Z < 2.32)$ could be transferred to either

$P(Z > -1.4) - P(Z > 2.32)$ for Table B users

or

$P(Z < 2.32) - P(Z < -1.4)$ for Table A users

However, she interpreted the format $P(-1.4 < Z < 2.32)$ as the area 'smaller than 2.32 and bigger than –1.4', so she transferred it to

$P(Z < 2.32) - P(Z > -1.4)$.

Similarly, she transferred $P(-2.41 < Z < -0.03)$ in Q1.6 to

$P(Z < -0.03) - P(Z > -2.41)$.

Arcavi (2003) argues that the "idea of betweenness" (2003, p.221) can be performed and corrected in the graphical representations if

> visualization… serves to adjust our 'wrong' intuitions and harmonize them with the opaque and 'icy' correctness of the symbolic argument (2003, p.222).

In addition, S7 mis-transferred $P(Z > -1.4)$ to $P(Z > 1.4)$ in Q1.5 and transferred $P(Z > -2.41)$ to $P(Z > 2.41)$ in Q1.6. The missing minus signs in these two questions indicated that she had difficulties with questions in such a format. S5 did not consider the area meaning and transferred Q1.6 to

$$P(Z > -2.41) - P(Z > -0.03) = 1 - P(Z > 2.41) - 1 - P(Z > 0.03)$$

because he regarded transferring as "how I calculated them" (S5, 1st May 09).

Although S6 solved Q1.1 correctly, he did not use the graph properly. S6's mark in the graph above Table A and Table B (Figure 11.1. and Figure 11.2.) indicated that he might have a problem locating a number on the number line without scales, or that he might not know the width of the normal curve. Using the attached graph with the left-tailed shadow area also made it difficult for S6 to perform or interpret a new shadow area clearly.

11.3.2 Graph users' and graph non-users' performances in Q3.1-Q3.2

In Q1.1-Q1.6, the two graph users and half of the graph non-users performed well in the transferring steps. However, both kinds of participants had difficulties finding $t_{0.99}(23)$ in Q3.2.

All participants clearly understood that this question concerned t distribution, and they used the table of t distribution accordingly. When they noticed that the table of t distribution did not include 0.99, but included 0.01, they attempted to transfer the symbolic question $t_{0.99}(23)$ including 0.99 to another symbolic form including 0.01. However, most participants did not consider the area meaning of 0.99, and their approach to transferring the question seemed to be influenced by the transferring principles of using tables of normal distribution in previous exercise questions Q1.1-Q1.6.

Although S3 and S6 drew graphs, their graphs (see Figure 11.3. and Figure 11.4.) did not properly portray the 0.99 as a big right-tail shadow area (compared with the small left-tailed white area 0.01) below the t curve. S3's simple graph only included a curve above an empty number line with a small right-tail labelled α. She did not consider how much space 0.99 should occupy in the right-tail, so her graph could not enable her to obtain the border of the shadow area in the left-hand side.

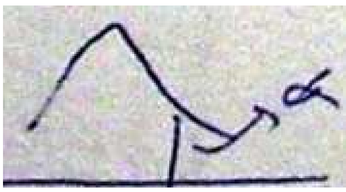


Figure 11.3. S3's graph in Q3.2.

S6 directly used the graph above the table of t distribution (see Figure 11.4.) when solving Q3.2. The shadow area in the attached graph was a small right-tail bordered on the right-hand side of 0, and S6 also did not consider how much space the area 0.99 should occupy and whether the graph he adopted matched the question or not.

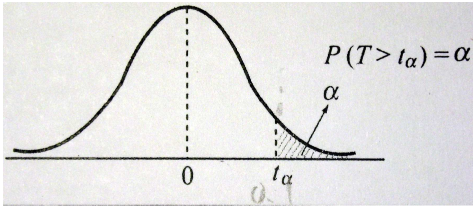


Figure 11.4. S6's graph in Q3.2.

Gray and Doritou (2006) find that some students may "exaggerate or underestimate" (p.73) magnitudes in the number line. In this study, I found that the two graph users had problems in deciding the size of the shadow area below the t curve. Although they correctly presented the shadow area on the right-hand side, they could not present the area in proper size. Therefore, their graphs did not match the question and could not help them to think about the question.

11.3.3 Graph users' and graph non-users' performances in Q15

Q15 was the only exercise question explicitly asking the participants to find z score(s). When looking for the z score of 90% confidence interval in the table of normal distribution without recalling it, S3 actively drew a graph with a big right-hand side shadow area and noted 0.01 (a result of mis-subtraction of 1 – 0.9) below the left-tail (see Figure 11.5.), but she did not answer the z score of 90% confidence interval because she could not find the exact value 0.01 inside Table B.

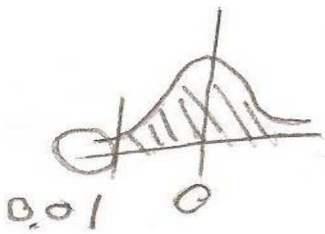


Figure 11.5. S3's graph of 90% confidence interval.

When looking for the z score of 95% confidence interval, S3 explained her approach as "there is no 0.95 in Table B, so I use 1 - 0.95. They are mutually matched (i.e.

symmetrical)" (S3, 29th Apr 09). She took the middle value 1.645 of 1.64 and 1.65, which correspond to 0.0505 and 0.4995 respectively, as the z score of 95% confidence interval. However, 1.645 was the z score of 90% confidence interval. According to S3's graph, we can see that she did not consider confidence interval as a middle chunk.

S5, S9 and S10 did not draw graphs in this question, and they found 1.96, the z score of 95% confidence interval, corresponding to 0.025 in Table B correctly and quickly. S1 also used the correct method, but he made a miscalculation in the beginning. S6, S7 and S8 could not find the value, so they attempted to recall it, but they did not recall the correct value. When it was suggested that they draw graphs, S6 drew a new graph of normal curve with two vertical lines on both sides of zero and regarded the middle chunk as the 95% confidence interval (see Figure 11.6.). His graph led him to notice 0.025 as a half of the other area 0.05 for both tails, and he found 1.96 corresponding to 0.025 as the z score of 95% confidence interval.

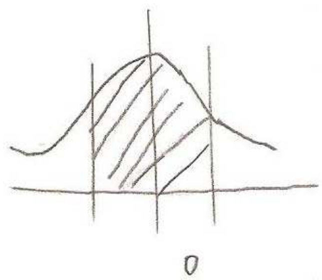


Figure 11.6. S6's graph of 95% confidence interval.

S7 drew a vertical line in the left-hand side of zero and a vertical line in the right-hand side of zero on the graph above Table B (see Figure 11.7.), and she explained that 95% is "from this [i.e. the vertical line on the left-hand side] to that [i.e. the vertical line on the right-hand side]" (S7, 4th May 09). However, S7 did not continue with her graph, but rather attempted to find the recalled z score 1.96 inside Table B. S2, S4 and S8 did not find the correct probability and did not have an answer to Q15.

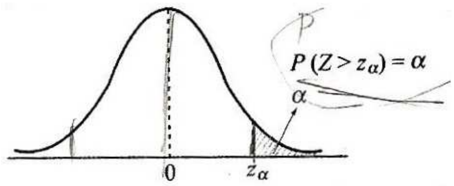


Figure 11.7. S7's graph of 95% confidence interval.

11.3.4 Graph users' and graph non-users' performances in Q1.a-Q1.d

When solving Q1.a-Q1.d in the beginning of the second interview, only S3 and S6 drew graphs in Q1.a, which appeared to be difficult for most participants. As mentioned in Chapter 7, most participants only used one table of normal distribution and transferring principles in the first interview, but they shifted to using Table A when the question involved 'smaller than' or to using Table B when the question involved 'bigger than' in the second interview.

Q1.a was difficult for the participants for two reasons. First, some participants did not interpret the given value in Q1.a and Q1.c as a probability because they were not familiar with such a question format. Students who relied totally on memorising principles frequently wrote inverse symbols such as 'bigger than' and 'smaller than', and/or 'negative' and 'positive', which was normally unseen in the transferring process. Some participants simply found a value visible inside or outside the table without considering its meaning. Some of them even switched the location of the symbols, so they unintentionally changed the meaning of the question and lost balance of the equation, such as S8 who changed

$P(Z < c) = 0.67$ in which 'c' was a z score

to

P(Z < 0.33) = c in which 'c' was a probability.

Second, even though some participants correctly interpreted the given value 0.67 as a probability, 0.67 was only included inside Table A (smaller than) and was not included inside Table B (bigger than), so the answer to Q1.a (bigger than) could not be directly found in Table A or Table B.

S3 represented the probability 0.67 as a big right-tailed area bordered on the left-hand side of zero in her graph (see Figure 11.8.) in Q1.a, but her graph did not help her to consider that 'a' on the left-hand side of zero was a negative value (see Appendix 7.2). She did not mark any symbol or number on the number line, and her answer of 'a' remained as a positive value. This example revealed S3's possible difficulty at interpreting the number on the number line without scales.

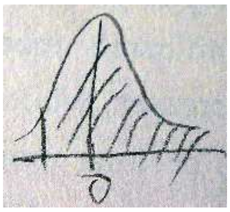


Figure 11.8. S3's graph in Q1.a.

S6 noticed that the given probability 0.67 was not inside Table B, so he used Table A (smaller than), but he was also clear that Q1.a was about 'bigger than'. He explained that

> This question is about bigger than, and I use the table [A] of smaller than, so I find 0.44 and add a negative symbol to get -0.44" (S6, 8th Jun 09).

That is, S6 found P(Z < 0.44) = 0.67 and applied the principle of P(Z < 0.44) = P(Z > -0.44). He also drew a graph (see Figure 11.9.) to represent the symmetric relationship between 'a' and '–a'.

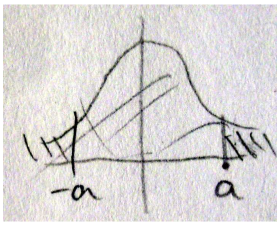


Figure 11.9. S6's graph in Q1.a.

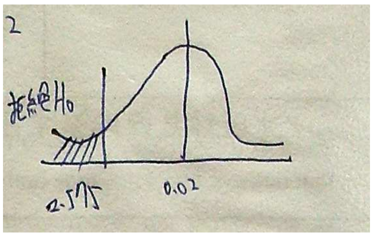
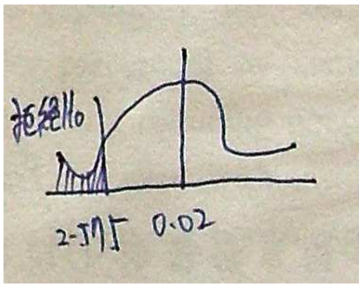
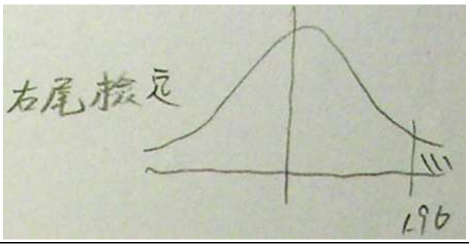
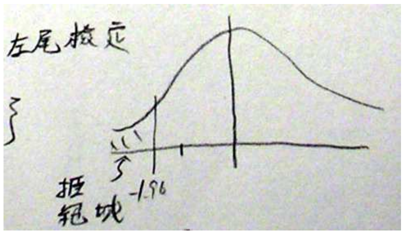
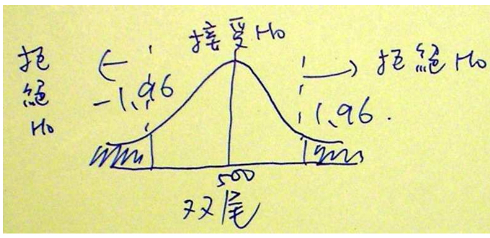
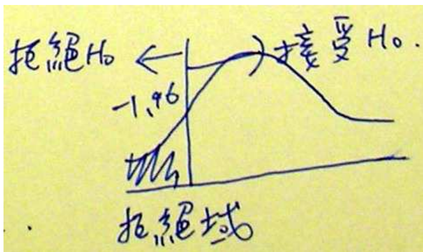
The 'a' in S6's graph represented the unknown value 'a', and the small shadow area on the right-hand side represented the given area 0.67. Tukey (1977) argues that a graph should "force us to notice what we never expected to see" (1977, p. vi). However, S6's right-tailed area was not big enough to inform him that 'a' was negative and located on the left-hand side of 0.

11.3.5 Graph users' and graph non-users' performances in Q6, Q7, Q9 and Q10

Three participants, S5, S6 and S7, drew graphs of the normal curve with shadow area(s) as critical region(s) in both Q6 and Q7 in the second interview (see Appendix 10.2.). S3 answered Q6 without a graph and did not answer Q7, but she drew graphs in both Q9 and Q10 (see Appendix 6.19.). However, S5 also attended the third interview, but he did not draw any graph in Q9 and Q10.

As mentioned in Chapter 10, many participants did not make correct hypotheses. However, the shadow area(s) in S5's, S6's and S7's graphs (see Table 11.2.) matched their alternative hypotheses. Since S7 was the only participant who assumed Q6 was a two-tailed test, her graph in Q6 was also the only correct graph with two tails. Although

S7 drew correct graphs in Q6 and Q7 and used correct critical regions in Q6, she did not obtain the correct critical value in Q7 because she obtained the critical value of Q6 and Q7 by recalling the z score of 95% confidence interval without considering how the one-tail shadow area influenced the z score. S3 appeared to decide the direction of the shadow area according to the test statistic Z_0 obtained by calculation in Q9 and Q10.

	Q6 (Q9 for S3)	Q7 (Q10 for S3)
S5	$H_0 : \mu = 500$ $H_1 : \mu < 500$ 	$H_0 : \mu \geq 500$ $H_1 : \mu < 500$ 
S6	$H_0 : \mu \leq 500$ $H_1 : \mu > 500$ right-tailed testing 	$H_0 : \mu \geq 500$ $H_1 : \mu < 500$ left-tailed testing 
S7	$H_0 : \mu = 500$ $H_1 : \mu < 500$ or $H_1 : \mu \neq 500$ two-tailed testing 	$H_0 : \mu \geq 500$ $H_1 : \mu < 500$ 
S3	$H_0 : \mu = 500$ $H_1 : \mu < 500$ $Z_0 = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \dfrac{500 - 498.2}{10 / \sqrt{100}} = 1.8$	$H_0 : \mu > 500$ $H_1 : \mu < 500$ $Z_0 = \dfrac{500 - 498.2}{10 / \sqrt{100}} = 1.8$

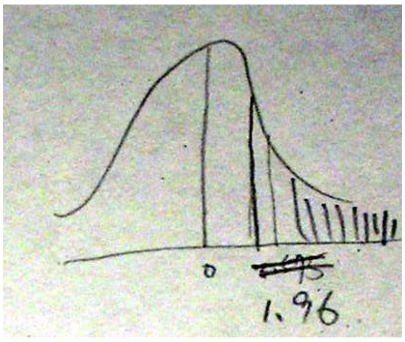
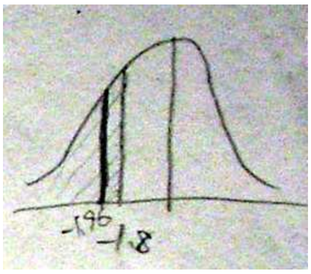
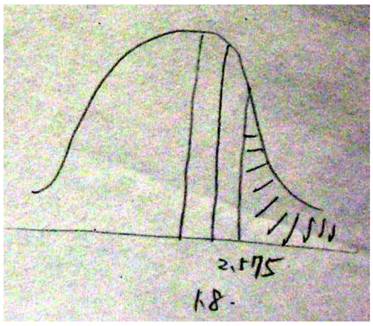
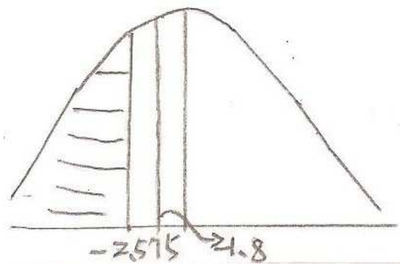
	90 95 $\rightarrow$ 0.05 99 $\rightarrow$ 0.1 $Z_{0.05} = \cancel{2.575}$ 1.96  Revised $Z_0 = \dfrac{498.2-500}{10/\sqrt{100}}$ = -1.8 	$Z_{0.05} = 2.575$  Revised $Z_0 = \dfrac{498.2-500}{10/\sqrt{100}}$ = -1.8 

Table 11.2. Four participants' hypotheses and graphs.

11.3.6 Comparison of performances of graph users and graph non-users

As mentioned previously, the two graph users (S3 and S6) and four graph non-users performed well in transferring steps in Q1.1-Q.1.6 and Q3.1-Q3.2. However, the other graph non-users often made incorrect transferring or did not transfer the equations even if transferring was necessary (see Table 11.3.).

	Graph user	Graph non-user
Performed better in transferring in Q1.1-Q1.6 and Q3.1-Q3.2	S3, S6	S5, S7, S9, S10
Performed worse in transferring in Q1.1-Q1.6 and Q3.2-Q3.2		S1*, S2, S4, S8

Table 11.3. Graph users' and graph non-users' performance in transferring in Q1.1-Q1.6 and Q3.1-Q3.2.

**S1's performance might be influenced by his use of Table B, which he was not familiar with.*

***Details of the participants' solving and transferring steps in Q1.1-Q1.6 and Q3.1-Q3.2 are listed in Appendix 7.1. and Appendix 8.1. respectively.*

The participants' performance in Q1.1-Q1.6 and Q3.1-Q3.2 was discussed in Section 7.3.1.1 and Chapter 8. In the following Table 11.4, I list whether the participants made correct or wrong transferring steps.

Participants	Transferring steps in	Q1.1	Q1.2	Q1.3	Q1.4	Q1.5	Q1.6	Q3.1	Q3.2
Graph Users	S3			C		CC	CCC		
	S6	C		C		CC	CCC		W
Graph Non-users	S1								
	S2								
	S4								
	S5					CC	CCC		W
	S7	C	C	C		WCW	W~~WW~~ CW		
	S8	W	W	W					W
	S9	~~W~~ C	W	C		CC	CCC		
	S10		C	C		C	C		W

Table 11.4. Graph users' and graph non-users' performance in transferring in Q1.1-Q1.6 and Q3.1-Q3.2.

**Transferring steps are considered correct (C) or wrong (W) only when they were written in the participants' solution. Blank means no transferring was taken or written in the solutions, no matter whether it was unnecessary or missed.*

***The ~~W~~ means that the wrong transferring step was deleted. The C or W in the second row means the participant's second trial.*

In Q1.1-Q1.6, the two graph users S3 and S6 had one hundred percent correct rates in all transferring processes, whereas graph non-users only had about a 65% correct rate (21/32) in the completed transferring chances. This instance provided evidence that graph users who successfully linked graphical representation and symbolic representation performed better in transferring steps than graph non-users who relied totally on symbolic manipulation. Other studies, such as that of Lowrie (2001), have similarly shown that students who predominantly use visual methods in mathematics problem solving tend to perform better.

However, graph users and graph non-users had problems with transferring Q3.2. Moreover, the graph users' graphs in Q15, Q1.a-Q1.d, Q6, Q7, Q9 and Q10 did not match the condition of the questions. For example, they did not represent a probability bigger than 0.5 as an area occupying more than a half below the curve, nor did they visualise confidence interval as a middle chunk. Therefore, the incorrect graph did not lead the graph users to obtain the correct answers.

11.4 The participants' graphical representations of normal distribution

When I examined the eight questions in the preparatory paper and mid-term examination in the second week of data collection, I found that six questions involved the formula or idea of standardisation (normalisation) (see Section 5.1). Standardisation, which involves the prerequisite knowledge of distribution, mean and standard deviation, appeared to be a dominant concept in the participants' statistics course. Therefore, I asked the participants to explain their understanding in words and in graphs in the first clinical interview, and most of them interpreted standardisation only as a formula (see Appendix 6.4. - Appendix 6.7.). When I analysed their responses, I sought to understand how they viewed the idea of distribution and how it relates to mean and standard deviation or variance.

Many of the recent studies in statistical education use qualitative methods to explore students' thinking and reasoning and how they develop. These studies find that many students have "relatively unsophisticated understandings of the concepts of mean and standard" (Garfield and Ben-Zvi, 2008, p.30). Therefore, I designed Q3 in the second interview to ask the participants to draw normal curves according to different means and standard deviations (variance):

> Q3. Present $X \sim N(20, 20^2)$, $Y \sim N(10, 10^2)$ and $\overline{X} - \overline{Y} \sim N(10, 5^2)$ by visualisation in a column.

S2 did not draw any graph, and S1's, S3's, S4's and S5's drawings in Q3 are listed and compared in Table 11.5.

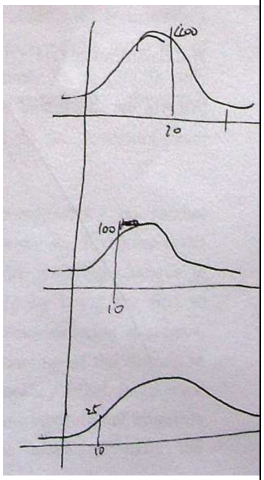
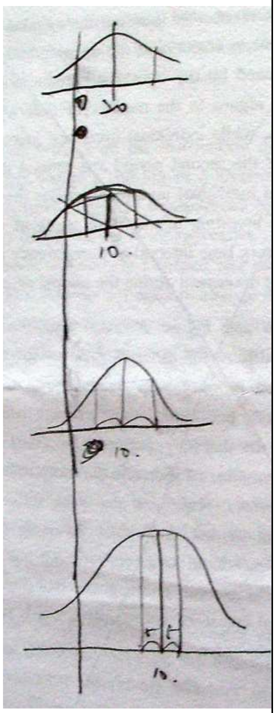
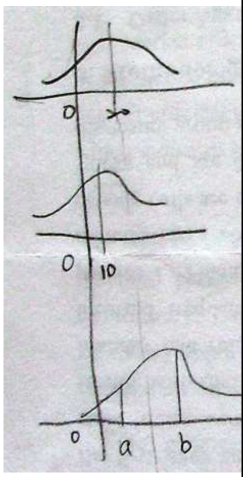
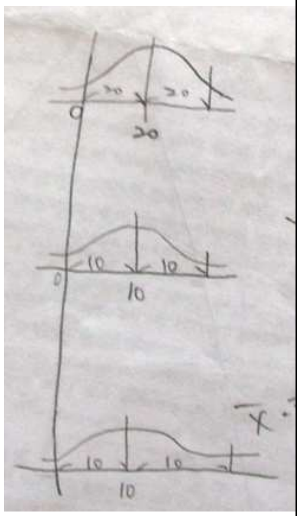
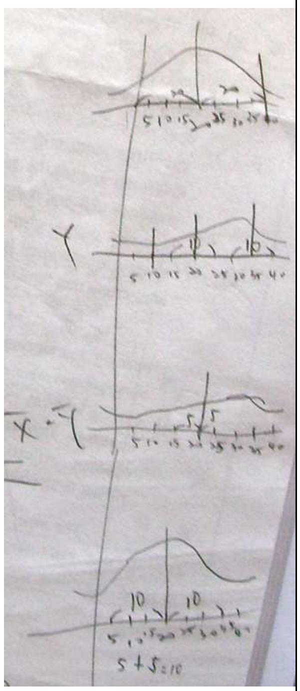
S1	S3	S4	S5
			 [S5 redrew.] 

Table 11.5. S1's, S3's, S4's and S5's drawings in Q3 in the second interview.

**S3 deleted the second curve. The second curve of S5's second trial was replaced by the fourth curve.*

S1 drew the three normal curves and marked variance 400, 100 and 25 above the three curves respectively. He explained that the third curve was flatter and less sharp than the second curve because the variance was smaller. In addition, it appeared that he did not determine the centre of the curve in terms of the mean.

S3, S4 and S5 did not know how to start drawing the curve of normal distribution, so I drew $X \sim N(20, 20^2)$ as an example, and then asked them to draw the other two curves below. S3 drew curves of $Y \sim N(10,10^2)$ and $\overline{X} - \overline{Y} \sim N(10,5^2)$ below my curve of $X \sim N(20, 20^2)$. S3's curves of $Y \sim N(10,10^2)$ and $\overline{X} - \overline{Y} \sim N(10,5^2)$ were in the third and fourth places respectively. She explained the curve of $Y \sim N(10,10^2)$ as "this is 10 and that is also 10" (the distance from the centre on the left-hand side and on the right-hand side) and the curve of $\overline{X} - \overline{Y} \sim N(10,5^2)$ as "the centre is 10, this distance is 5 and the other distance is 5" (S3, 3rd Jun 09).

There are several findings in S3's graphical representation. First, she represented the length of 10 as longer than the length of 5. Second, she did not consider the location on the x-axis since the location 10 of the centre of the curve of $Y \sim N(10,10^2)$ was the same as the location 20 of the centre of the curve of $X \sim N(20, 20^2)$, but different from the location 10 of the centre of the curve of $\overline{X} - \overline{Y} \sim N(10,5^2)$. Morever, the left point with distance 10 away from the centre of $Y \sim N(10,10^2)$ was not located on 0. Third, S3 did not decide the spread or width of the curve according to the standard deviation or variance since the second curve with standard deviation 10 was as wide as the first curve with standard deviation 20 but narrower than the third curve with standard deviation 5.

S4 also drew the curves of $Y \sim N(10,10^2)$ and $\overline{X} - \overline{Y} \sim N(10,5^2)$ below my curve of $X \sim N(20, 20^2)$. She considered the location of number on the x-axis, but she did not understand what the 10 and 5^2 in $\overline{X} - \overline{Y} \sim N(10,5^2)$ meant and how they influenced the location and spread of the curve. She simply wrote them (a was 10 and b was 25) below the x-axis.

S5 drew the curve of $Y \sim N(10,10^2)$ and $\overline{X} - \overline{Y} \sim N(10,5^2)$ centred at 10, but they were in the same location and shape as my curve of $X \sim N(20,20^2)$. He also seemed not to notice that the standard deviation in the third curve was not 10 but 5.

In order to explore whether he determined the centre and spread of the normal curve according to the mean and the standard deviation, and whether he considered the location and length in the graphical representation, I drew another graph of the curve of $X \sim N(20,20^2)$ with graduation 5, 10, 15, 20, 25, 30, 35 and 40 below the x-axis and asked him to draw the curves of $Y \sim N(10,10^2)$ and $\overline{X} - \overline{Y} \sim N(10,5^2)$ on the x-axis with such axis again. S5 drew the curve of $Y \sim N(10,10^2)$ in the second place and the curve of $\overline{X} - \overline{Y} \sim N(10,5^2)$ in the third space, but he drew another curve of $Y \sim N(10,10^2)$ in the fourth place to replace the second curve.

S5 did not determine the centre and the spread according to the mean and standard deviation. Instead, he developed his own principles to draw the curves. His principles were explained in our dialogue:

> I: Why are all curves centred on 20? In the last curve of $Y \sim N(10,10^2)$, what does the distance 10 in the left-hand side mean?
> S5 (pointing at the first 10 of $Y \sim N(10,10^2)$): This
> I: How about the distance 10 in the right-hand side?
> S5: The same.
> I: Then what does the centre 20 mean?
> S5: It is 5 away from 5 to 10, and it is also 5 away from 15 to 20, and 5 plus 5 equals 10.

I: What is the distance from 5 to 20? How do you deal with the sector from 10 to 15?
S5: I just added the two sectors (which means from 5 to 10 and from 15 to 20)
I (pointing at the first 10 of $Y \sim N(10,10^2)$): Ok. How do you deal with the 10?
S5: 5 plus 5 equals 10.
I: What does 10 represent?
S5: I think 10 is the distance. If I calculate and get the distance 10, then I decide the points.
(S5, 5th Jun 09)

S5 seemed to view the mean in the symbolic representation of normal distribution as the distance from the centre, and he determined the location of the centre of the curve according to the result of addition. Although S5 marked the length of standard deviation on the x-axis, he did not decide the shape of the curves based on the standard deviation.

I tested Q3 to S1-S5, but some of them did not know how to start drawing the graph, and some of them did not consider the location or length on the x-axis or shape of the curve in their graphical representation. Therefore, I gave the graph of normal curve of $X \sim N(20,20^2)$ with graduation 0, 10, 20, 30 and 40 on the x-axis as an example (see Figure 11.10), and I asked S6-S10 to draw the normal curve of $Y \sim N(10,10^2)$ and $\overline{X}-\overline{Y} \sim N(10,5^2)$ on such x-axis below the curve of $X \sim N(20,20^2)$:

Q3. $X \sim N(20,20^2)$ is shown on top. Please draw $Y \sim N(10,10^2)$ and $\overline{X}-\overline{Y} \sim N(10,5^2)$ in the second and third places.

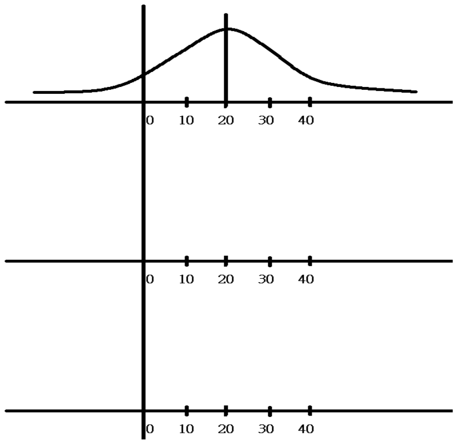
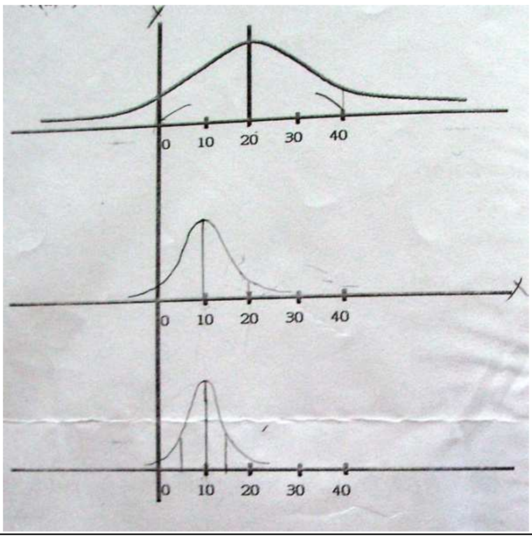
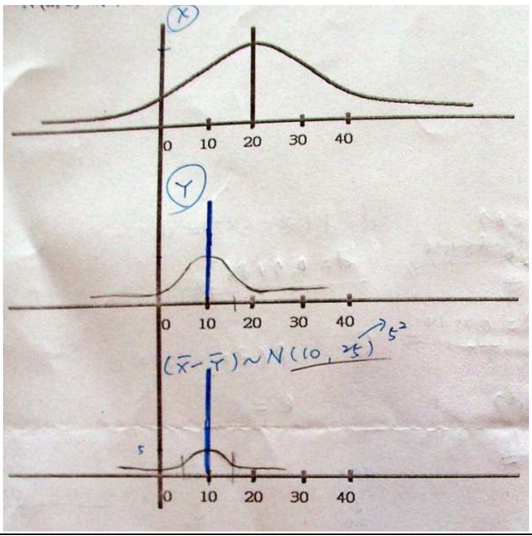


Figure 11.10. Example and places for drawing of Q3 to S6-S10.

S10 did not like graphs and he did not draw. The other four participants' drawings are listed and compared in Table 11.6.

S6	**S7**
	

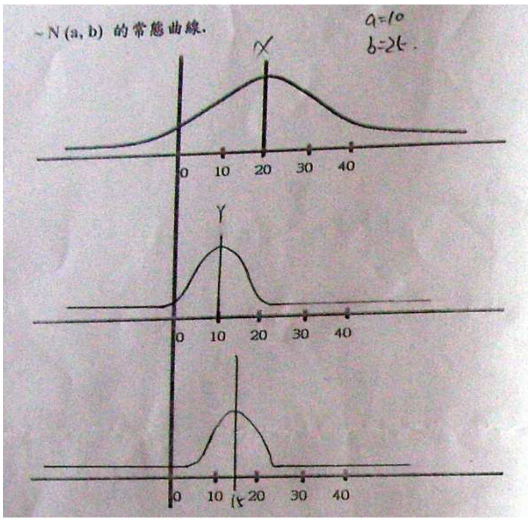
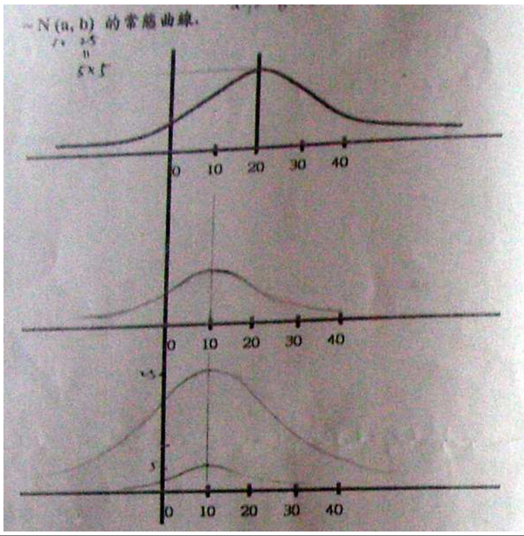
S8	S9
	

Table 11.6. S6's, S7's, S8's and S9's drawings in Q3 in the second interview.

S8 drew a vertical line to confirm the location of the top point, but she did not determine the centre of the curve according to the mean. She used different principles to draw the two curves. For instance, she determined the centre of the curve of $\overline{X}-\overline{Y} \sim N(10,5^2)$ according to the subtraction of 10 and 25(5^2). She believed that her curve of $Y \sim N(10,10^2)$ was sharper than the curve of $X \sim N(20,20^2)$ because 10 was smaller than 20. She decided the shape of the curve according to the mean.

Both of S9's curves were centred at the means, but he determined the heights rather than the width of the curves according to the standard deviations. He noticed that there seemed to be a square of width 20 and height 20 in the graph of $X \sim N(20,20^2)$, and that there was also a square of width 10 and height 10 in the graph of $Y \sim N(10,10^2)$. In his graph of $\overline{X}-\overline{Y} \sim N(10,5^2)$, the width of the rectangle remained 10 and the height became 25, so the curve became taller and sharper. As soon as he noticed that 25 was not the standard deviation but the variance, he got the standard deviation 5 and drew another shorter and flatter curve.

S7 found that the centre of the drawn curve of $X \sim N(20, 20^2)$ was determined by the mean and that the height of the curve "seems to correspond to 20". Therefore, she drew the curve of $Y \sim N(10, 10^2)$ and the curve of $\overline{X} - \overline{Y} \sim N(10, 5^2)$ according to this rule. Her curve of $Y \sim N(10, 10^2)$ was narrower and lower than the curve of $X \sim N(20, 20^2)$, and her curve of $\overline{X} - \overline{Y} \sim N(10, 5^2)$ was even narrower and lower. She talked about the height of the curve, but she did not mention the width. Her curves were very similar to S9's curves.

S6 determined the centre of the curve according to the mean and the widths of the curve according to the standard deviation, so the curve of $Y \sim N(10, 10^2)$ was narrower than the curve of $X \sim N(20, 20^2)$, and the curve of $\overline{X} - \overline{Y} \sim N(10, 5^2)$ was even narrower.

Apart from S2 and S10, the participants drew normal curves according to the symbolic representation of normal distribution. Most of them did not draw the correct curves, and their drawings revealed reasons for their difficulties. First of all, some participants could not recognise what the elements in the symbolic representation meant. For example, some of them could not distinguish standard deviation or variance in this format.

Second, some participants did not determine the centre of the normal curve according to the mean. Third, some participants did not determine the spread of the normal curve according to the standard deviation. Rather, they used the standard deviation to determine the height of the normal curve.

Fourth, some participants used inconsistent principles to draw the two curves. For instance, S8 determined the centre of the curve of $Y \sim N(10, 10^2)$ according to the front value 10 and the centre of the curve of $\overline{X} - \overline{Y} \sim N(10, 5^2)$ according to the difference of 10 and 25.

Fifth, most participants were not concerned with the length and location on the x-axis, or with the information the factors expressed in graphical representation. Such a situation was similar to some participants who were not concerned with the location of the z score in the graph of normal curve in other exercise questions. Their inattention to the information in the graphs, such as locations and lengths, might explain why they were unable to perform, compare or take advantage of areas which represented probabilities in the exercise questions.

11.5 Summary

In this chapter, I first introduced the participants' interpretation of and attitude towards visualisation and how visualisation relates to this study. I also explained the benefits of using graphs and why students refused to use graphs in this study. Then I compared the graph users' and graphs non-users' performances in the five types of exercise questions and analysed the problematic graphs. According to the participants' solutions and graphs in some exercise questions, such as Q3.2 and Q1.a, we can see that many participants never considered whether the unknowns in Q3.2 and Q1.a were positive or negative and simply treated them as positive numbers. It is found in much literature that, for many students, "number means exclusively the natural numbers" (Verschaffel *et al.*, 2006, p.52). Although some participants used graphs because they believed them useful, graphs were useful only when they matched the structural aspects of the problems (Corter and Zahner, 2007). Participants' reluctance or inability to correctly visualise the questions prevented them from being conscious of the fact that the requested numbers were negative. This was also found in Eisenberg and Dreyfus (1991):

> …students… do not know how to take advantage of those [graphs] provided to them or even those they have drawn themselves (Eisenberg and Dreyfus, 1991, p.33).

At the end of this chapter, I examined the participants' drawing of normal curves according to the symbolic representation and found that some participants did not determine the centre and spread of the normal curve according to the mean and standard deviation respectively. The participants did not link the elements of symbolic representation and of graphical representation correctly. Moreover, they did not pay attention to some of the important components (i.e. location and length) in the graph. These phenomena reflected and explained the participants' difficulty in taking advantage of graphs in other exercise questions in this study.

CHAPTER 12

REVIEW OF DATA ANALYSIS

The main part of this study consists of three clinical interviews with the ten participating students. The exercise questions tested the participants' learning in four parts of statistical inference which they learnt in the statistics course: using the tables of normal distribution, using the table of t distribution, looking for the z score of 90%, 95% and 99% confidence intervals, and carrying out the whole processes of hypothesis testing including looking for the critical value(s) and critical region(s). Models of solving steps in the different types of question were developed to analyse and compare the participants' solutions step by step.

I will briefly review the exercise questions in the order in which they occurred in Section 12.1, and then I will review the participants' responses to these questions in Section 12.2.

12.1 Review of the exercise questions

At the beginning of the first interview, the participants were asked to find the probability value in the tables of normal distribution according to relation sign(s) and z score(s) in Q1.1-Q1.6, which were expressed in symbolic form, such as $P(Z < 1.22)$, $P(Z > -3.1)$ and $P(-1.4 < Z < 2.32)$. The questions asked for the probability of a standard normal random variable Z being smaller or bigger than a given z score, or the probability of Z being between two given z scores. Both Table A and Table B were taught by the teacher in class and could be used by each participant in the first interview, except for S1 and S2, who were only given Table B and did not ask for Table A. Eight participants preferred Table B while S1 and S4 preferred Table A, and only S4 used Table A in the first interview.

In the following Q3.1 and Q3.2, the participants were asked to find the t score according to the given probability and degree of freedom in the table of t distribution. The two questions were also expressed in a symbolic format ($t_{0.005}(9)$ and $t_{0.99}(23)$ respectively) but they looked different from the format of Q1.1-Q1.6. The format of Q3.1 and Q3.2 was the same as the format of $t_{\alpha}(v)$ in the participants' textbook and in the teacher's writing. Q3.2 was difficult for most participants, and only two participants applied the correct approach. One explanation of the difficulty might be because the given probability was not included in the table of t distribution. Q3.1 and Q3.2 were the only two questions of t distribution in the entire study.

At the end of the first interview, the participants were asked in Q15 to find the frequently used z score of 90%, 95% or 99% confidence interval rather than simply recalling them. That is, they should find z score in the table of normal distribution according to the given probability, but the relation sign was not given or implied, and the given probability 90%, 95% or 99% was not the probability value they should find in the table of normal distribution. Therefore, the participants needed to construct the symbolic equation by themselves, or alternatively, they needed to do so without writing the equation and by considering and comparing the meaning of the given numbers and elements in the tables to find the answer. These questions were also problematic for most participants.

According to the analysis of participants' difficulties in Q3.2 and Q15 in the first interview, Q1.a-Q1.d were designed and tested at the beginning of the second interview in order to examine whether the participants were able to distinguish and recognise the elements of symbolic questions and how they connected the elements in symbolic form with the elements in the tables of normal distribution. The two questions Q1.a and Q1.c in new formats tested whether and how the participants found the z score according to the given relation sign and probability. Q1.a was in a format which did not fit the format of Table A or Table B unless it was transferred, and it was difficult for most participants since only three of them applied correct approaches. Q1.b and Q1.d were in the basic format of Q1.4 and Q1.2 respectively, but some participants seemed to be confused by the similar symbolic formats and could not distinguish Q1.a, Q1.b, Q1.c and Q1.d.

At the end of the second interview, the participants solved two questions of hypothesis testing. Q6 was a two-tailed test and Q7 was a left-tailed test. The questions explored how the participants made a null hypothesis and an alternative hypothesis, which approach they used, how they found the critical region and how they made decisions. All participants used the test statistic approach, so the critical value(s) was/were z scores which could be found in the tables of normal distribution. S5, S6 and S7 used graphs in both questions. S1-S5 did not perform well in the second interview, so they attended the third interview to solve the two questions again as Q9 and Q10. S5 did not draw any graphs, but S3 drew graphs in Q9 and Q10.

12.2 Review of the participants' responses

Careless mistakes such as mis-reading the question or mis-calculating equations, which occur in many mathematics questions, were also made by the participants in the exercise questions of statistical inference. The mistakes in the order of appearance in the solutions, especially of using the tables of distribution to find a value, consisted of mis-reading, mis-transcribing, mis-transferring, mis-targeting, mis-interpreting and mis-calculation. Mistakes simply caused by carelessness were identified in the early stage of data analysis and excluded from further investigation in this study, and other kinds of problems in solving the questions were explored and analysed. According to the analysis of the participants' responses to the exercise questions, most of the problems seemed to be caused by the participants' false interpretation of the symbolic language and meaningless manipulation of the symbolic representations.

According to Aberdein (2007), "mathematical luck" (2007, p.13) may occur in mathematical processes in which some errors made in different stages compensate for and cancel each other, so a correct answer cannot necessarily indicate that the correct approach was applied. Such mathematical luck was also observed in some participants' solutions. Moreover, an incorrect answer may be obtained by the correct approach with careless mistakes. Therefore, whether the answer is correct or incorrect is not an ideal indicator of understanding how to solve the question. Thus, this study examined whether the correct solutions were made (correct solution means that correct methods and steps were applied, and the correct answer would be obtained if there was no careless mistake).

The participants' solutions of using the tables of normal distribution and t distribution to solve each type of exercise question are summarised below.

12.2.1 Using the tables of normal distribution (Q1.1-Q1.6)

Some participants had problems when using their preferred Table A or Table B of normal distribution with or without transferring to find the probability of Z being smaller or bigger than a given value or the probability of Z being between two values. These problems occurred regardless of whether the answer could be directly found in the preferred table or not. The participants' main problems were inattention to the relation sign and minus sign in the equation, manipulation of symbols which did not maintain the equivalence of the equation and false or even reverse interpretation of the elements inside and outside the tables. S3 and S6 used graphs (taking advantage of the attached graphs in the textbook or drawing new ones by themselves) in some questions smoothly.

12.2.2 Using the table of t distribution (Q3.1-Q3.2)

Most participants had difficulties when they used the table of t distribution to find the t score according to the given probability and degree of freedom in symbolic form $t_{0.99}(23)$ in which the probability 0.99 was not included inside the table. They did not consider the area meaning of the given probability and attempted to transfer the question in a way similar to the transferring principles they used in normal distribution, so the equivalence of their equations was not maintained. S3 and S6 also used graphs but their graphs were not correct and did not help.

12.2.3 Using the tables of normal distribution to find values in different formats (Q1.a-Q1.d)

Although the participants chose Table A or Table B of normal distribution according to the relation sign of the question, most of them had difficulties in the new format of looking for the z score according to the given relation sign and probability, especially when the answer could not be directly found in the table. Some participants could not recognise whether it was the probability or the z score they should look for according to the symbolic equation, even after comparing the question to the format which they often solved and used. S3 and S6 also used graphs to aid their thinking, but there were some problems with their graphs.

12.2.4 Looking for the z score of 90%, 95% and 99% confidence intervals (Q15)

Most participants had difficulties finding the z scores of 90%, 95% and 99% confidence interval in the tables of normal distribution without recalling the memorised ones. One explanation of their difficulties was that the participants did not consider the meaning of confidence interval (with two tails) and attempted to find the probabilities 1 – 0.90, 1 – 0.95 and 1 – 0.99 in the table of normal distribution. The participants were also asked to recall the three z scores, and nearly one-third (5/16) of their memorised z scores were incorrect yet similar to the correct ones. Three participants used graphs, but only S6's graphs proved helpful.

12.2.5 Carrying out the whole processes of hypothesis testing (Q6, Q7, Q9 and Q10)

Most participants had difficulties carrying out the whole hypothesis testing procedure. Unlike previous exercise questions, which simply asked the participants to find the answer in the tables of normal distribution or t distribution, the two questions of hypothesis testing required the participants to decide whether the assumption was correct or not, so the participants had to first make a null hypothesis and an alternative hypothesis and then decide whether the test was two-tailed, left-tailed or right-tailed according to the narrative question, find the critical value(s) and critical region(s), obtain the test statistic by using the formula and calculation, compare the test statistic and critical value and finally decide whether the assumption was correct or not. The participants used the test statistic approach, but very few of them completed all of the necessary steps. Each of the steps seemed to be problematic for some participants, and I will explain the participants' problems in each step as follows.

First, some participants were unaware of the rule requiring them to make the alternative hypothesis match the sample. Only one participant made correct two-tailed hypotheses according to the narrative condition ‘equal to’ in Q6, while half of the participants made correct left-tailed hypotheses according to the condition ‘smaller than’ in Q7. S1-S5 were retested in Q9 and Q10, but still only one of the participants correctly made two-tailed hypotheses in Q9. Their hypotheses in Q6, Q7, Q9 and Q10 are listed in Table 12.1. and Table 12.2.

	Q6 (equal to)	Q7 (smaller than)
correct	$H_0: \mu = 500, H_1: \mu \neq 500$ (two-tailed)	$H_0: \mu \geq 500, H_1: \mu < 500$ (left-tailed)
S1	$H_0 = 500, H_1 < 490$	$H_0 < 500, H_1 > 500$
S2		
S3	$H_0: \mu \geq 500, H_1: \mu < 500$	
S4		
S5	$H_0: \mu = 500, H_1: \mu < 500$	$H_0: \mu \geq 500, H_1: \mu < 500$
S6	$H_0: \mu \leq 500, H_1: \mu > 500$ right-tailed testing	$H_0: \mu \geq 500, H_1: \mu < 500$ left-tailed testing
S7	$H_0: \mu = 500, H_1: \mu < 500$ or $H_1: \mu \neq 500$ two-tailed testing	$H_0: \mu \geq 500, H_1: \mu < 500$
S8	$H_0: \mu = 500, H_1: \mu < 500$	$H_0: \mu \geq 500, H_1: \mu < 500$
S9	$H_0: \mu \geq 500, H_1: \mu < 500$	$H_0: \mu \geq 500, H_1: \mu < 500$
S10	$H_0: \mu > 500, H_1: \mu < 500$	

Table 12.1. S1-S10's hypotheses in Q6 and Q7 in the second interview.

**Hypotheses made by participants were parallel. They are listed in this way for convenience of comparison.*

	Q6 (equal to)	Q7 (smaller than)
Correct	$H_0: \mu = 500, H_1: \mu \neq 500$ (two-tailed)	$H_0: \mu \geq 500, H_1: \mu < 500$ (left-tailed)
S1	$H_0: \mu = 500, H_1: \mu \neq 500$	$H_0: \mu > 500, H_1: \mu \leq 500$
S2		
S3	$H_0: \mu = 500, H_1: \mu < 500$	$H_0: \mu > 500, H_1: \mu < 500$
S4	$H_0: \mu \leq 100, H_1: \mu > 100$	$H_0: \mu \geq 100, H_1: \mu < 100$
S5	$H_0: \mu = 498.2, H_1: \mu \neq 498.2$	$H_0: \mu \geq 500, H_1: \mu < 500$

Table 12.2. S1-S5's hypotheses in Q9 and Q10 in the third interview.

**Hypotheses made by participants were parallel. They are listed in this way for* *convenience of comparison.*

Whether the test is two-tailed, right-tailed or left-tailed must be determined in the beginning because it influences not only the absolute value of the critical value, but also whether the critical value is positive or negative. Although nine participants mistakenly regarded Q6 as a one-tailed test, many of them still obtained correct critical values. On the other hand, many participants correctly regarded Q7 as a left-tailed test, but none of them obtained the correct critical region. When the test is two-tailed, critical regions on both sides need to be determined, but some participants did not consider the negative critical region. Moreover, five participants made left-tailed testing in Q6, but their critical values were positive. Also, S5 made two-tailed hypotheses in Q9, but he compared the negative test statistic to the positive critical value. Some participants did not consider the existence of negative critical values.

Some participants did not correctly interpret the meaning of the numbers given in the two narrative questions and mistakenly substituted them in the hypotheses, such as using the digits of sample size or sample mean rather than the assumed population mean in the hypotheses. This did not influence their critical region, but it did influence the test statistic.

Some participants used the correct absolute value of critical value Z = 1.96, $Z_{0.025} = 1.96$, or 1.96 with α = 0.05 in Q6, although they wrongly made a left-tailed alternative hypothesis. The participants' correct critical values in Q6 were found to be a lucky coincidence by comparing the same critical values used by the participants in Q7 with different conditions. The critical values in Q6 ($\pm z_{\frac{0.05}{2}}$) were 1.96 and -1.96 because Q6 was a two-tailed test, but the critical value in Q7 ($-z_{0.05}$) was not -1.96 but -1.645

since Q7 was a left-tailed test. Some others recalled the z score 1.645 or 2.575 as the critical value in both Q6 and Q7. In other words, most participants used the memorised z scores of confidence interval as the critical value in both questions without considering the type of test.

Third, five participants had critical regions in Q6, and two of them made their critical region in the right-tailed even if they made left-tailed alternative hypotheses. Only one participant, S7, made correct two-tailed hypotheses in Q6; three participants, including S7, found the correct absolute value of the critical value, but only S7 found the correct negative critical regions. On the other hand, five participants made correct left-tailed hypotheses in Q7, but no participant found the correct critical value or critical region, and four participants used the same critical value as that in Q6.

Four participants used graphs, but their graphs did not prove particularly helpful. The participants' solutions to the two questions indicated that the influence of the type of hypothesis testing was not considered by the participants when looking for critical values and critical regions. One explanation for this might be that the participants did not distinguish between the confidence interval and critical region. Most participants, except for S1 and S3, used the same critical value in both questions even if some of them interpreted Q6 as a two-tailed test and Q7 as a left-tailed test.

CHAPTER 13

RESEARCH FINDINGS AND DISCUSSION

As stated in Chapter 1, this study aims to explore the following research questions:

1. How do the students use the tables of distribution to solve questions in statistical inference, and what difficulties do they have?
2. How does visualisation help students use the tables of distribution, and what difficulties do they have?
3. Why do some students not use visualisation to help in using the tables of distribution, and how do the students who use visualisation perform in comparison with those who do not use visualisation?

In Section 13.1 to Section 13.3, I will illustrate and discuss the findings in response to each of the research questions.

13.1. Research question 1:

> How do the students use the tables of distribution to solve questions in statistical inference, and what difficulties do they have?

By deconstructing the participants' solutions, I have found that the participants had several kinds of problems in solving the questions, and their problems appeared to be related to the use and interpretation of symbols. The participants' difficulty in one question or even in one step might be caused by a combination of multiple problems. I have distinguished six distinct problematic types of symbol use/interpretation and listed them in Table 13.1. As mentioned in Chapter 7, mis-reading and mis-calculation will not be discussed here.

Types of problem / Problems with the symbols	Using the tables of normal distribution (Q1.1-Q1.6)	Using the table of t distribution (Q3.1-Q3.2)	Using the tables of normal distribution to find values in different formats (Q1.a-Q1.d)	Looking for the z score of 90%, 95% and 99% confidence intervals (Q15)	Looking for the critical value(s) and critical region(s) (Q6, Q7, Q9 and Q10)
Miss the minus sign	**2**	**3**			**6**
Find the number in the table directly and meaninglessly	**5**		**3**		
Manipulate the number mistakenly or meaninglessly	**12**	**7**	**5**	**2**	**11**
Interpret the symbol and number in different places of the equation reversely			**4**		
Interpret the meaning of the memorised number inside and outside the table reversely				**1**	
Interpret the number inside or outside the table mistakenly	**3**	**2**	**2**	**2**	

Table 13.1. The participants' problematic use of symbolic representation.

**Problems that occurred in the participants' revised solutions were also counted, but if the same problem happened in one solution it was only counted once.*

***The problems in the question of hypothesis testing in this table are limited to the stage* *of looking for the critical value(s) and critical region(s).*

The first difficulty occurred in two situations. One situation was that two participants did not notice the existence of the minus sign in the question. The other situation occurred in the following questions, where negative z score or t score was not considered as a potential answer by most participants. In other words, negative numbers were excluded in their considerations. Problems involving negative numbers in equations have also been noticed by other researchers. For example, Vlassis (2002) finds that some students subtracted 5 in both sides of the equation $6x - 5 = 7$, because of "the minus sign in front of 5 not having been taken into consideration", and difficulty arises from the "inability… to imagine the sign in front of the expression" (Vlassis, 2002, p.349).

The second difficulty happened to four participants who avoided transferring steps and directly found a value corresponding to the given value in the table of distribution as the answer without considering what the given/requested value was and whether the equation with the given value and the found value could be balanced. In other words, the participants solved the questions without any meaning-making. Although the "detachment of meaning… coupled with a global "gestalt" view" (Arcavi, 1994, p.31) of symbolic representations is needed for quick and efficient manipulations, Arcavi (1994, 2005) argues that the students with symbol sense who read through the symbolic representations towards meaning can make reasonable solutions since symbolic representation always expresses the whole scenario and carries its explanation. However, some participants simply found a value corresponding to the given value in the tables of distribution as the answer without considering the meaning of the given value, requested value and the question.

The third type of problem appeared to be significant since it occurred in all kinds of exercise questions, and all of the participants had such a problem at least once. When the given value was included in the tables, some participants found it directly (i.e. the second problem), whereas some others made “automatic manipulation” (Arcavi, 2005, p.43), such as using the transferring principles correctly or incorrectly.

In this study, automatic manipulations were found most frequently when the given probability was not included in the tables. One common situation was that the participants transferred the equation by manipulating the relations sign, the minus sign and/or the given number without considering what the symbols and values meant and whether the equivalence of the equation was maintained. Another situation happened to two participants who did not transfer but subtracted two probabilities corresponding to the two given z scores. Yet another situation occurred in the second question of t distribution, in which the given probability was not included in the table of t distribution. Half of the participants seemed to be led by their “met-befores” (Tall, 2004, p.286) of transferring, so they transferred the symbolic form in a very similar way to what they had done in the previous and typical question of looking for a probability according to a given z score. Therefore, difficulties occurred when the students’ concept aligned poorly with the standard meaning of the concepts (Patrick *et al.*, 2010). Each of these participants did not possess enough symbol sense of

> [the] realization of the need to *check for the symbol meanings during the implementation of a procedure*, the solution of a problem, or, during the inspection of a result, and the comparison and contrasting of those meanings with one’s own intuitions about the expected outcome (Arcavi, 2005, p.43).

Students who have adequate symbol sense (Fey, 1990) feel it necessary to check for symbol meaning, but some participants without symbol sense might make "typical and counterproductive" (Thompson *et al.*, 2010, p.1) manipulations in their solutions when the given value was a probability which was not included inside the tables, such as looking for the z scores of confidence interval or critical region of two-tailed testing.

The fourth kind of problem was observed with four participants when they used the tables of normal distribution to find a value according to the new question formats with a given probability (i.e. Q1.a and Q1.c). These participants still found probability in the direction of outside-in as their answer, and they were not able to distinguish the meaning of what was given and what was requested in the symbolic representation even if they had a chance to compare the new question formats with previous question formats.

The fifth problem only occurred once, when one participant tried to find a recalled z score inside the table. Her solution indicated that she memorised the coupled numbers (95% and 1.96) without understanding their meanings. Although this participant manipulated the symbolic equations well, she lacked symbol sense to identify the likely form of the relationship between the elements inside and outside the table of normal distribution and the meaning of the elements by scanning the table and her graph (Fey, 1990).

The sixth problem was due to the special manner of expressing the probabilities and the z scores in the tables. For instance, the probability .0202 in the table was interpreted as 0.202; the z score 0.44 expressed as .4 in the left-side column and .04 on top of the table was interpreted as 0.404; and .0005 on top of the table of t distribution was interpreted as 0.005. Such a situation led the participants to obtain wrong values in the tables.

In this section, we have seen that all of the participants had problematic use of symbolic expressions to varying degrees. Many of them focused on procedures instead of giving meanings to the equations. Therefore, they simply relied on symbolic manipulation and the tables of distribution to find the answers. The tables of distribution link to the structures of graphs with the function to "display data" and "organize information" (Friel *et al.*, 2001, pp. 127-128), but the notions and the features of distributions draw on the behaviours exhibited in graphs of data, and statistical knowledge should be related conceptually to the pictures (Wild, 2006). In the next section, I will discuss how visual presentations related to the participants' learning, question solving and difficulties in statistical inference.

13.2. Research question 2:

> How does visualisation help students use the tables of distribution, and what difficulties do they have?

According to the analysis of the participants' responses to the exercise questions in the interviews (see Chapter 7-11), most of the participants' problems were connected to the symbolic representations of the exercise questions and of the formulas representing the relationship between the z scores or t scores and probabilities in the tables of distribution, including $P(Z < z) = A$, $P(Z > z_\alpha) = \alpha$ and $P(T > t_\alpha) = \alpha$. I will explore this issue in this section. The influence and difficulty of the graphs used by four participants to help solve the questions will also be discussed.

Two tables of normal distribution and one table of t distribution were used in this study. Each of the three tables was attached in the participants' textbook with a graph. To be more specific, there is a graph with a left-tailed normal curve above Table A (see Figure

4.14), a graph with a right-tailed normal curve above Table B (see Figure 4.15) and a graph with a right-tailed t curve above the table of t distribution (see Figure 11.1).

In this study, most exercise questions involving the use of the table of distribution were presented in symbolic form, such as Q1.1-Q1.6 and Q3.1-Q3.2 in the first interview, and Q1.a-Q1.d in the second interview; others were word problems, such as Q15 in the first interview (see Appendix 6.3.), Q6 and Q7 in the second interview (see Appendix 6.11.) and Q9 and Q10 in the third interview (see Appendix 6.18.). In those questions, the participants needed to find the probability, z score or t score according to the information given. However, the participants were not explicitly told what type of value they should search for, except for Q15. Some participants used only the tables of distribution to solve the questions, while some others used both the tables of distribution and the graphs to solve the questions.

Graphs helped the users in several ways. When the relation sign and the z score were given, such as in Q1.1-Q1.6, Q1.b and Q1.d, locating the z score in the graph enabled the user to see it on the left-hand side or right-hand side of zero on the number line. Visualising the relation in the question enabled the user to see the direction of the shadow area. Therefore, the user could have an idea of what the shadow area is like, such as a big/small left/right tail, or a big/small middle chunk. Also, comparing the drawn graph and the graph of the table helped the graph user decide whether the table was suitable or unsuitable for directly finding the value.

On the other hand, when the probability and the relation sign were given (i.e. Q1.a and Q1.c), drawing the shadow area in the correct direction and proper size helped to raise awareness of the location of the z score. However, in Q3.2, Q15, Q6, Q7, Q9 and Q10, the relation sign was not mentioned, and the given probability was not necessarily the probability to be found in the tables of t distribution or normal distribution. Understanding the questions and drawing graphs required prior knowledge of where the given probability is, and this caused difficulty in drawing correct graphs. To be more specific, the 0.99 in Q3 was a large area on the right-hand side; the 95% confidence interval in Q15 was not in left-tail, right-tail or two-tails, but rather the middle chunk below the normal curve; and $\alpha = 0.05$ should be placed in the left-tail in Q7 as a whole, whereas the shadow areas in both tails in Q6 were not 0.05 but 0.025 (half of 0.05). These questions were found to be very difficult for most of the participants. Some participants attempted to create meaning by visualising the questions, but it also proved difficult for them to draw graphs in these questions, and they might not have drawn correct graphs. With a wrong graph or the absence of a graph altogether, the participants often found the t score or z scores corresponding to the probability obtained by meaningless symbolic manipulation.

Most participants did not use graphs, and only participants S3 and S6 used or drew graphs in all kinds of questions, which were basic use of the tables of normal distribution and t distribution, using the tables of normal distribution to find the z scores of 90%, 95% and 99% confidence intervals, and carrying out hypotheses testing. Another participant, S7, used a graph to find z scores of confidence interval and solve hypotheses testing, and a fourth participant, S5, drew graphs in the question of hypothesis testing. In Table 13.2., we can see when graphs were actively used or drawn by participants in a few questions.

	No.	Type of question	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
The first interview	Q1.1	Using tables of normal distribution to find probability (symbolic question)						CC				
	Q1.2											
	Q1.3							CC				
	Q1.4											
	Q1.5				CC			CC				
	Q1.6				CC			CC				
	Q3.1	Using table of t distribution to find t score (symbolic question)										
	Q3.2				IN			II				
	Q15	Finding the z score of confidence interval (narrative question)			II			CC	CIR			
The second interview	Q1.a	Using tables of normal distribution to find probability or z score (symbolic question)			CIR			IIR				
	Q1.b											
	Q1.c											
	Q1.d											
	Q6	Hypotheses testing (narrative question)					IIR	II	CIR			
	Q7						IIR	CIR	CIR			
The third interview	Q9				~~II~~ CIR							
	Q10				~~II~~ CIR							

Table 13.2. The participants' graphs in their solutions.

**Blank means no graph was involved. Slanted line means the question was not asked. S3's first graph in Q9 and first graph in Q10 were abandoned.*

There were five kinds of graphs in the participants' solutions. Some of the graphs used or drawn by the participants were correct, included sufficient necessary information and led the participants to correct solutions (such graphs are represented by CC), while some other graphs were incorrect, included at least one problem and led the participants to incorrect solutions (II) or influenced the participant's thinking and she had no solution (IN). However, there were also some graphs irrelevant to the solutions, no matter whether the graphs were correct (CIR) or incorrect (IIR).

Most of the CC graphs appeared in the questions involving the use of the tables of normal distribution to find a probability, and the II, IN, CIR and IIR graphs were used or drawn in almost all the questions using the tables of normal distribution to find the z scores or using the table of t distribution to find the t score.

Three significant problems were found in these graphs. First, the participants were unable to represent a right-tailed area which was bigger than 0.5 below the t curve or z curve in Q3.2 and Q1.a, and they represented it as a small right-tailed area bordered in the right-hand side of zero, which implied a positive t score or z score. This might be because one graph user did not interpret 0.99 in the symbolic form $t_{0.99}(23)$ as a probability which can be represented as an area, or because she had difficulty expressing area 0.99 below the t curve in her graph (see Figure 13.1.). Therefore, she did not consider a negative answer in Q3.2.

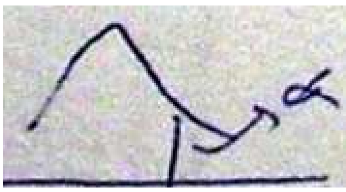


Figure 13.1. S3's graph in Q3.2.

Another graph user marked the border on the original border t_α of the attached graph (see Figure 13.2.) above the table of t distribution in his textbook. This might have led him to regard the value at the border as positive, thereby preventing him from considering a negative answer.

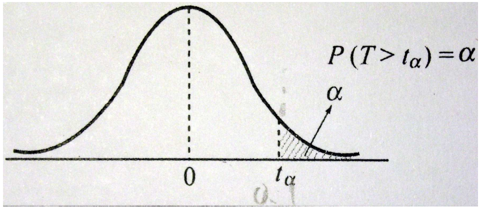


Figure 13.2. S6's graph in Q3.2.

Second, even if the probability was correctly represented as a big right-tailed area bordered on the left-hand side of zero in Q1.a (see Figure 13.3.), the graph did not help the user to obtain a negative z score of 'a'. There was not any negative number marked on the number line, and her answer of 'a' remained as positive. Therefore, her graph was correct but irrelevant to her symbolic manipulation to obtain 'a'.

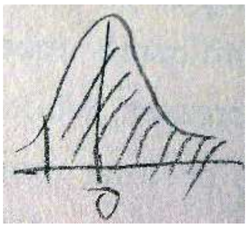


Figure 13.3. S3's graph in Q1.a.

The three examples above indicate the participants' difficulty of scaling in the graphs, including scaling a big area and scaling a number on the left-hand side of zero as negative. Moreover, according to the participants' graphs in Q3 in the second interview, they had difficulty locating numbers on the number line.

Figure 13.1., Figure 13.2. and Figure 13.3. also reveal the importance of interpreting elements in symbolic form, expressing them in visual form and interpreting information brought by visual representation. As stressed by Tukey (1977),

…the greatest value of a picture is when it forces us to notice what we never expected to see (Tukey, 1977, p. vi).

Third, in this study, the focus of the solution of Q6, Q7, Q9 and Q10 was on looking for the critical region(s) with positive and/or negative critical value(s). The participants' correct or incorrect graphs were determined by their hypotheses, but most of their critical values were irrelevant to the graphs with left-tail, right-tail or two tails and instead came from their memorised z scores of confidence intervals. Such behaviour was shown in some graph users' writings and graphs with the same z score in both questions.

In this section, I have reviewed and discussed how the participants used visual representation in learning and solving questions in statistical inference and what difficulties they had when using graphs. Although visualisation serves as a "tool for meaning-making" (Breen, 1997, p.97) and both symbolic representation and visualisation are at the core of understanding (Duval, 1999), graphs were refused by most participants in the study. Only four of them with adequate symbol sense took advantage of graphs to visualise what they saw from the symbolic representations in a few questions, and some of their graphs were not correct. Therefore, many of the participants' solutions showed a lack of meaning, and this resulted in difficulties in solving equations (Lima and Tall, 2006b). In the next section, I will explain why a visual approach was not applied by the participants.

13.3. Research question 3:

Why do some students not use visualisation to help in using the tables of distribution, and how do the students who use visualisation perform in comparison with those who do not use visualisation?

S2, S4, and S8 felt that the graphs of normal distribution and t distribution in statistical inference were too difficult to learn and use. S9 and S10 relied totally on formulas to solve questions and did not attempt to use any graph. S10 said that "… if I see a graph, I will just ignore it" (S10, 4th May 09).

S7 did not like to use graphs, nor did she feel the necessity to use the graphs of normal distribution and t distribution. She believed that "you can solve questions without drawing graphs… drawing graphs is troublesome and unnecessary" (S7, 4th May 09).

As mentioned previously, the two graph users and four graph non-users performed well in transferring steps in Q1.1-Q.1.6 and Q3.1-Q3.2, whereas the other four graph non-users usually made incorrect transferring or did not make the necessary transferring steps (see Table 11.3.).

In Q1.1-Q1.6, the two graph users S3 and S6 had one hundred percent correct rates in all transferring processes, whereas graph non-users only had about 65% correct rate in the thirty-two completed transferring chances. The results indicated that graph users who correctly linked graphical representation and symbolic representation performed better in transferring steps than graph non-users. However, many of the graph users' graphs in Q3.2, Q15, Q1.a-Q1.d, Q6, Q7, Q9 and Q10 did not correctly represent the condition of the questions, and the incorrect graph did not lead the graph users to obtain the correct answers.

In short, by investigating the four research questions in the study, we can see the benefits of using graphs in learning and solving questions of statistical inference, and it is evident that graph users performed better in the transferring steps than graph non-users. However, we also hear several graph non-users' voices: the graphs of tables of distribution are too difficult to learn and use, or they are troublesome and unnecessary (see Section 11.2.1 and Section 11.3.1). Some participants' graphs did not represent the correct condition, and the graphs did not lead them to obtain correct solutions because of the participants' problematic interpretation of symbolic representations.

13.4 Discussion

Based on my previous teaching experience in statistics courses, this study begins by investigating the participants' use of the tables of normal distribution and t distribution. When the question format matched the formats of the chosen table of normal distribution, the users could directly find the answers in the table; when the question format did not match the format of the chosen table, transferring steps needed to be applied to transfer the question to the suitable format before using the tables. The teacher asked the students to remember the transferring principles; therefore, the rote-learnt meaningless procedures (in this study, to transfer the equations) might be easily misunderstood and misused by students (Lima and Tall, 2006a).

Lima and Tall (2006a) explain that some students see equations as a calculation with numbers and unknowns to get the solution, but they only possess procedural ways of thinking without conceptual meanings of symbols. Their finding supports Kieran's (1981) argument that some students see the equals sign as a "do something symbol" (Kieran, 1981, cited in Lima and Tall, 2006a, p.1), which leads them to perform an action, a calculation, instead of representing equivalence between the two sides of an equation. Lima and Tall (2006b) stress that keeping a balance of equations requires not only performing the same operations on both sides, but also giving meaning to symbols, but the idea of equivalence or balance was not noticed by students. Cortes and Pfaff (2000) find that some students transfer equations by meaningless procedures without justification in order to get a result.

Three participants' inattention to the relation symbol and negative sign was influenced by their choices of the table of normal distribution. Two Table B users treated the 'smaller than' symbols as 'bigger than' and treated the negative value as a positive value, whereas the Table A user treated the two questions of 'bigger than' as questions of 'smaller than'. The three participants thus avoided using the transferring principles which they felt difficult, but they still accidentally changed the format of the equation in their writings because they wanted to directly look for the given value in the chosen table of normal distribution and take the corresponding value in the table as the answer, without considering how to make the equation match the format of the chosen table. Such an approach was explained by the Table A user as follows: "I don't understand the transferring formula… [so] I only use the table" (S4, 29th Apr 09).

Most participants had difficulties in the questions with negative answers (that the answers were negative was not revealed implicitly or explicitly by the questions). Some students perhaps believed that a minus sign represented a negative number and read -x as minus x or negative x, so they would regard -x as a negative number. In this way, the 'a' in Q1.a did not reveal that the unknown value might not be positive, so the participants might not actively consider it as negative. Thus, only three participants obtained negative answers in this question, while six of the other participants had positive answers. This situation was also found in Q3.2. Legutko (2008) has another explanation for this: the students unconsciously ignore the minus sign because they have "more experience in operations on positive numbers than on negative numbers" (2008, p.146).

In the questions in different formats, participants also had problems interpreting the numbers or symbols. Freudenthal (1983) stresses that letters in mathematics mean something as well as symbols, and what they mean must be made perfectly clear.

Like most arithmetic questions, the required answers or results of Q1.1-Q1.6 and Q3.1-Q3.2 are on the right-hand side of the equal sign. However, some participants did not notice the new format of Q1.a, whose answer was on the left-hand side of the equal sign, and they still applied the typical and familiar method of solving questions in the format of Q1.b, whose required answer was on the right-hand side. Mellin-Olsen (1987) argues that being familiar with such a new language is a basic requirement for thinking, because sometimes the main problem of under-achievers is not lack of knowledge but discontinuity between the knowledge forms.

Garfield and Ben-Zvi (2008) argue that students' reasoning and understanding of statistical inference might be caused by traditional probability theory-based explanations in formal statistical language, and it is better to introduce the idea of inference informally at the beginning of the course by simulation or re-sampling methods to avoid over-mathematising. With adequate knowledge of what the elements in the equations are and the relationship between them, practising the functional language may help overcome difficulties (Mellin-Olsen, 1987). Freudenthal (1983) emphasises that

> …the rules for applying structuring devices (as well as transformation rules) should be learned by *using* them – consciousness is required for fighting *misuse* (1983, p.472).

Booth (1981) also finds that there are two coexisting systems of mathematics: a “formal taught system” and a “system of child-methods [in order to] solve mathematical problems within a ‘human-sense’ framework” (1981, p.29). Students normally do not use taught algorithms unless their own methods have failed (Hart *et al.*, 1981, cited in Mellin-Olsen, 1987, p.92). Therefore, offering students the opportunity to practise questions in different formats enables them to not only discover their difficulties but also recognise the different formats of questions of using the tables of normal distribution and revise their approaches.

A common situation occurred in the questions which were more difficult for the participants. In Q3.2, some participants did not realise the meaning of the given value 0.99, which was not included in the table of t distribution, so they used subtraction to obtain a value included in the table and transferred the equation by using some principles which were similar to the transferring principles of normal distribution. In Q15 and Q1.a, some participants did not consider the meaning of the given value and had difficulties obtaining the correct value to be found in the tables of normal distribution, so they could not proceed or treated it as a question in a different format. In Q6, Q7, Q9 and Q10, the participants obtained the critical values by substituting the z scores of confidence interval, which were memorised or found in the tables, without considering the meaning and influence of two-tails and one-tail. Their solutions were deeply influenced by their “met-befores” (Tall, 2004, p. 286) of previous questions. Although such earlier and old experiences could be fundamental knowledge when constructing new ideas, they might have a negative effect and cause conflicts (Tall, 2002). Correct approaches for previous questions were not suitable but were nevertheless used by the participants in the exercise questions, so most of their solutions were incorrect.

In many exercise questions of this study, some participants carried out procedures in order to get a result, but the procedures were confused and misused. Without considering the meanings of the questions, the given value and the requested value presented in symbolic form, some participants who relied totally on memorising principles wrote ‘bigger than’ or ‘smaller than’ reversely, or they missed the negative sign, and some of them switched the location of the given value and the requested value. Therefore, they unintentionally changed the meaning of the question and lost balance of the equation. Some participants attempted to use an operational relationship between two numbers by finding a value corresponding to the given value in the table as the answer without realising or considering the relationship between the paired values, and some others meaninglessly used the rules to solve questions (Lima and Tall, 2006a, 2006b).

Wild (2006) explains that in the standardised histograms of continuous measurement data, proportions are represented by areas. This format involves probability density and its curve. Through standardisation, two or more data sets can be compared. Guzmán (2002) explains that the image is the matrix from which concept and method arise.

Stix (1994) explains that a multi-modal approach leads to a deeper and truer understanding of mathematics since it enables students to connect visual imageries, verbal knowledge and individual experiences (1994). Thus, a student with the flexibility to choose multiple solving methods is supposed to perform better than another who only uses a formula. One graph user felt it more difficult to memorise the formulas (i.e. transferring principles) without errors than to remember graphs, so she preferred to indirectly “remember the formula by remembering the graphs” (S3, 29th Apr 09). This is also mentioned in Arcavi (2003), as “the visual display of information enables us to ‘see’ the story … and possibly to remember it vividly” (2003, p.218). In order to take advantage of the graphs,

> one must go from the whole graph to some visual values that point to the characteristic features of the represented phenomenon or that correspond to a kind of equation and to some characteristic values within the equation (Duval, 1999, p.17).

The “whole story” (Arcavi, 2003, p.218) of the graphs of distribution can be told in the two-dimensional graph with shadow area(s) below the curve by indicating its broad features including types of curve, direction of shadow area, size of shadow area and location of z or t score(s). The ‘total area 1’ below the normal curve or the t curve was not explicitly noted in the graphs of normal distribution and t distribution, and this might be a reason why it was frequently misused or ignored by the participants (even the graph users). Lack of indication of total area 1 in the graphs made graph users unable to take advantage of the idea of ‘complementary’ of the shadow area(s) and the other white area. Such a situation was found in Q3.2, and Q15 and Q1.a.

Participants might not link the value inside the tables to the shadow area below the curves. It is found in Mundy’s (1984) study that students did not consider the integrals of positively valued functions in terms of an area under a curve (Mundy, 1984, cited in Eisenberg and Dreyfus, 1991).

Therefore, in Q1.5, one participant directly subtracted two found probability values corresponding to the given z scores without considering the minus sign in order to get the middle. Mundy (1984) explains that some students only have a mechanical understanding because they have not achieved a “visual understanding of basic underlying notions” (1984, cited in Eisenberg and Dreyfus, 1991, p.27).

Healy and Hoyles (1996) notice that students "rarely exploit the considerable potential of visual approaches to support meaningful learning" (1996, p.67). Clements (1984) observes that students, even those that are mathematically advanced, may avoid using visualisation (1984). Some participants who performed well in transferring refused to use graphs because they felt drawing graphs troublesome and unnecessary. Even if some participants attempted to take advantage of graphs, they had problems creating correct graphs. Therefore, the meaning of the questions could not be correctly represented, and the participants could not solve the questions with the right meanings.

Using the tables of normal distribution and t distribution is a basic and important skill in learning statistical inference, but there is little previous research about how students learn to use these tables and what difficulties they may have. Therefore, this study addresses this issue and wishes to fill the gap. Two participants actively used graphs in all kinds of exercise questions, and two other participants used graphs in some of the questions. The other six participants did not use graphical representation in the entire study.

Most participants who simply relied on symbolic manipulation were found to have significant difficulties when the requested values could not be found directly in the tables. On the other hand, the graph users performed better in some questions whose requested values could not be found directly in the tables, but they also had difficulty dealing with the questions with given probability bigger than 0.5. Siegler and Booth (2004) find that students may "exaggerate" (2004, p.429) or underestimate magnitudes (i.e. the length) in the number line, and the graph users in this study appeared to underestimate the area and were unable to represent an area bigger than 0.5 below the curve of distribution. More than one-half of the participants did not consider confidence interval as the middle chunk, and they could not distinguish confidence interval and critical region, so they did not notice how the two-tailed and one-tailed test influenced the critical value(s). Therefore, they could not obtain the correct probability for searching in the tables.

13.5 Conclusion

This study investigated year-two college students' learning in statistical inference with/without the use of visual methods. I began this study by presenting the four research questions, and I then collected and analysed the data, which led me to focus on three particular aspects of statistical inference: using the tables of distribution, looking for the z score of confidence interval and looking for the critical value according to the significant level. The second and third aspects are applications of the first aspect.

I designed exercise questions to focus on the three aspects, and I tested the questions with the participants in the clinical interviews in order to explore their solving strategies. Also, I developed a framework of solving steps by refining Kirsch's *et al*. (1998) process variables in order to analyse the participants' strategies in solving the exercise questions. The framework of solving steps allowed me to deconstruct the solution of using the tables of normal distribution to find z scores or probabilities and using the table of t distribution to find t scores. The deconstruction helped me notice each of the problematic steps, and I then analysed and explained why the problems occurred. Furthermore, the analysis provided evidence of the participants' six types of problems with the symbolic representations. With some adjustments, the framework can also help with deconstructing and analysing the steps of using the table of Chi-square distribution and the table of F distribution in future research.

This study revealed that most of the participants' difficulties in learning the three aspects were caused by the fact that the meaning and location of the elements of the equation, table and graph were not correctly linked and that area meaning was not absorbed in their learning of transferring. Some participants manipulated symbols without considering what their manipulations meant or without keeping the balance of the manipulated equations.

This study explores the benefits of using graphs of distribution in learning and solving questions of statistical inference and provides evidence that graph users performed better in the transferring steps than graph non-users did. This study also explores why some participants refused to use graphs. Moreover, some participants drew graphs, but their graphs did not lead them to obtain the correct solutions because they did not correctly construct the graph in terms of the elements of the table or equation, such as locating a given negative z score on the left-hand side of zero on the number line, or presenting a given probability as a shadow area below the curve in a proper size. Some participants, even the graph users, appeared not to consider and take advantage of the ideas of area (i.e. symmetry or complementary) in transferring steps.

In this chapter, I have reviewed the four research questions of the study and discussed how the study responds to these research questions. This study suggests that the elements of symbolic equation, table and graph must be linked, and the area meaning of the graph must be strengthened in the teaching and learning of using the tables of distribution. In the next chapter, I will discuss reflections and implications of this study.

CHAPTER 14

REFLECTIONS AND IMPLICATIONS

14.1 Reflections

Corbin and Strauss (2008) argue that one of the most important factors of a study is the quality of material being analysed, and I acknowledge that there are some shortcomings in my collected data. First, I had no experience of conducting clinical interviews before this study, so my skill of interviewing and ability of designing questions were perhaps imperfect. In order to enhance my ability of interviewing, I read articles about interview skills, attended training courses and seminars, and had discussions with my colleagues and supervisors about my interview questions and the processes beforehand.

However, some problematic situations still arose in the clinical interviews. For instance, in the first interview in the ninth week, I did not give both tables to the participants because I had observed that only Table B was used by the teacher in the eighth week. One of the participants normally used Table A, and he used the unfamiliar Table B to answer the questions. This might have caused some problems in his solutions. I did not record how each participant found each answer in the table, and I did not consider its consequence during the interviewing process. I found some answers of the participants difficult to analyse, and they did not remember how they had found the answer on previous occasions. Although I used a video camera to record how some participants found their answers in the table in the second interview, it was still impossible to record all of the group interviewees. Moreover, I did not ask the same questions of all participants, such as in Q15 in the first interview or Q3 in the second interview, and this caused some problems in comparing their solutions.

Second, the interviews took longer than the expected length originally agreed upon with the participants, so in some questions, they did not have enough time to think thoroughly, and I did not have time to break the ice for each student in each question. This might have prevented some participants from thinking as thoroughly as they would have liked, so the questions might not sufficiently explore the participants' real thoughts.

Third, as mentioned previously, only one participant was interviewed individually, and the other nine participants were interviewed in two groups of two and one group of five members respectively. There was one less capable interviewee in each group, each of whom did not talk much. In some questions, the less capable interviewees kept silent or said "I guess" or "I don't know" when I asked them to explain their solutions. In other responses, they seemed to repeat a peer's explanation without revealing their own thoughts. I noticed such situations when I transcribed and analysed the first interview. Therefore, in the second and third interviews, I checked each interviewee's solution after they finished solving and let the interviewees who had exhibited difficulties and had written less than their peers to talk about their thinking first (Brown and Liebling, 2005).

Fourth, I did not pay much attention to the students' interactions in the discussions, since each of them solved most of the questions on their own and their discussion was not supposed to influence the written answer. However, some participants followed others' discussions to find the z scores of confidence interval at the end of the first interview. Their answers were mixed with their peers' ideas, and they became confused. I noticed this situation and prevented it from happening again in the second and third interviews, usually by encouraging these students to respond first.

Fifth, although it is good to combine observation with interviews (Corbin and Strauss, 2008), I did not have enough data on student observation since my classroom observation was focused on the teacher, rather than on the ten participants who sat randomly within fifty students and who did not speak much in classes. Also, the teacher did not collect assignments from the students, so there was no material to collect except for the students' answer sheets for the mid-term and final examinations. However, the answers were given before the examinations, and many participants admitted that they had memorised the answers.

Last but not least, as introduced in Chapter 1, the study was inspired by my three-and-a-half years of experience in statistics classes, with nearly three hundred students. A large number of them had problems when using the tables of distribution. Typical questions of using the tables of distribution required the ability to find the corresponding values as the first step or as a whole, but I found that many users of the tables developed their own approaches. I attempted to search relevant literature, but there was little research concerning this issue, which led me to construct the present study. Based on my experience in statistics classes and my findings in the participants' statistics class, I designed exercise questions to explore the ten participating students' solving steps and problems of using the tables of distribution. I tried to analyse the solutions in an objective way, so I repeated comparing each participant's solution in each question to his/her solution in other questions, referred to their explanation in the interviews, and checked the video recordings if available in order to find the pattern of each student's solution, but I have to admit that some important leads might be missing and need to be uncovered due to the subjective nature of this approach. Also, the findings of the study cannot represent all students learning the tables of distribution, but

I believe that they reveal an underestimated learning difficulty in statistical inference. I hope that the study can inspire similar research in the future.

14.2 Implications

14.2.1 The importance of learning to use the tables of distribution

According to the results of the participants' performance in the exercise questions, we can see that some participants who did not use the table of normal distribution well also had difficulties finding the z score of confidence intervals or the critical value/region in hypothesis testing. That is, the ability to use the tables of distribution actually influenced the outcome of learning statistical inference.

To be more specific, three participants had difficulty solving the basic questions of using the table of normal distribution because of their inattention to the relation sign and the negative sign. They did not distinguish the different question formats, nor did they consider the meaning of the questions. Instead, they directly found the given value in the table of distribution. Thus, they gave up looking for the z score of 95% confidence interval because 0.95 was not included in the chosen Table B. The meaningless solving strategy also caused a participant to obtain a critical region on the right-hand side even though she had made a left-tailed alternative hypothesis. Because the three participants had difficulties using the tables of normal distribution, they could hardly solve any questions in the statistical inference. Similar situations had also been found in my statistics courses, an experience which led me to carry out this study and design this series of questions.

14.2.2 The video in my statistics course

In the analysis of the participants' solutions, we can see how symbolic language caused difficulties for the students. Questions in such symbolic presentation were used in the textbook and introduced by the teacher before the students realised their meaning. Learners are expected to read formulas with understanding, but formal substitution is one of the main difficulties in learning algebraic language (Freudenthal, 1983). Therefore, it is important for teachers to help students obtain symbolic meanings instead of relying completely on symbolic manipulation. In order to rectify such problems, teachers should ask the learner not to jump to the symbolic manipulation right away, but rather to make sense of the problem and the symbols, to draw a graph and to reason their implementation (Arcavi, 2005). Taking Q1.a for example, instead of saying the value corresponding to 0.67, it is better and more meaningful to say the probability of Z being bigger than 0.67. Without giving meaning to the question, such difficulties cannot be prevented. This difficulty may be improved upon by watching the video of normal distribution.

The video which was used in my statistics course and led me to construct this study was introduced in Chapter 1. According to the positive feedback from some of my students, this video more or less helped them understand the table of normal distribution. However, some information was left out of this video because these graphs and video were made within only a few days. The first thing I should have included is the formula representing the relationship between the z score and corresponding probability. It would also have been useful to explain transferring principles in the video. For example, when the red left-tail meant $P(Z < -2.00)$, I could explain by showing the formula that the white right-tail represented $P(Z > -2.00)$ and could be obtained by $1 – P(Z <-2.00)$.

This video was not played to the participants for three reasons. First, I was clear that my role was not a teacher but a researcher, and I should not teach the participants without the teacher's permission. Second, most participants only used Table B in the first interview, and watching my video about Table A might have caused some confusion. Last, I did not wish for the participants to compare the teacher's and my teaching. However, such a video may be used in other statistics classrooms to help students better understand the relationship between the elements of the table, elements of the graph and elements of the formula. With some revisions, a video about Table B can also be made.

14.2.3 Implications for examining students' understanding of the tables of distribution and further research

Duval (1999) and Arcavi (2003) find that many students can construct graphs according to the given equations but that they cannot discriminate equations when looking at Cartesian graphs because they often confine themselves to the irrelevant visual variables and details. The participants in my study did not have a chance to write down the equations according to the graph of normal distribution or t distribution in their classes or examinations. Although this was not tested in my interviews, it may be a good method to test the students' understanding of the relationship between the elements inside and outside the table and elements in the symbolic equation by asking them to construct the symbolic question according to the graph with z score on the x-axis and probability represented by a shadow area on the left or right-hand side.

Also, negative z scores and probabilities which are bigger than 0.5 were not included in Table B, and this caused many difficulties to my participants. I believe it is important for statistics teachers and researchers to notice this problem and help to remedy it.

There is a third table of normal distribution, and each value inside the table represents the probability of Z being between 0 and the corresponding z score (see Figure 14.1.). I did not discuss this table because it was not mentioned by the teacher or the participants in the entire process of data collection. However, this table was also used in some

Taiwanese textbooks, and it might therefore cause difficulties to students. How students use this table, what benefits this table has, what difficulty students may have and whether this table influences how students learn confidence interval and hypothesis testing deserve further investigation.

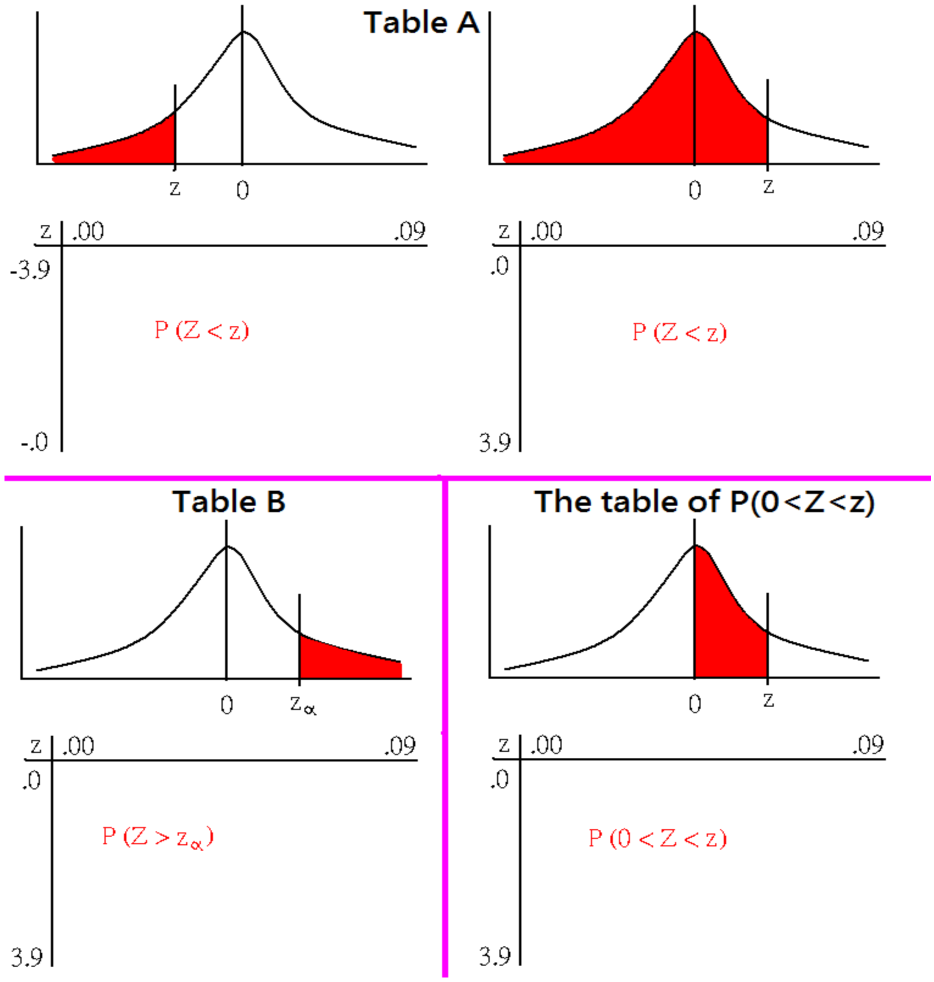


Figure 14.1. A brief comparison of Table A, Table B and the table of $P(0 < Z < z)$.

14.3 Conclusion

Although there were some limitations in my data, I analysed them and categorised the participants' solving behaviours, and I found a general problem within their solutions. Most of the wrong answers reflected a lack of meaning since graphs were considered unnecessary, and some participants avoided or openly refused to incorporate graphs. Using graphs was also problematic for some graph users. Such an issue has not been investigated in other studies, and I hope that this study can generate some ideas and interests. For example, the models of solving steps for each format of question enable teachers or researchers to realise which steps the students have problems in. In addition, videos of varying graph and table of normal distribution, or other computer programmes which can show the formula, graph and table can help to familiarise learners with the relationship between the elements.

REFERENCES

Aberdein, A. (2007). *Learning from Our Mistakes – but Especially from Our Fallacies and Howlers*. PMP 2007: Perspectives on Mathematical Practices. Vrije Universiteit Brussel, March 2007.

Anghileri, J. (2000). *Teaching Number Sense*. London: Continuum.

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics, *For the Learning of Mathematics,* **14** (3), 24-35.

Arcavi, A. (2003). The role of visual representations in the learning of mathematics, *Educational Studies in Mathematics*, **52** (3), 215-24.

Arcavi, A. (2005). Developing and using symbol sense in mathematics, *For the Learning of Mathematics,* **25** (2), 42-48.

Artigue, M., Batanero, C. & Kent, P. (2007). Mathematics thinking and learning at post-secondary level. In F.K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp.1011-1049). Greenwich, CT: Information Age Publishing Inc, & NCTM.

Batanero, C. (2004). Statistics education as a field for research and practice. *Proceedings of ICME-10*. Regular Lecture. Copenhague, 2005, International Commission for Mathematical Instruction. No pagination (hereafter, "n. p."). Online: http://scholar.google.co.uk/scholar_url?hl=en&q=http://www.ugr.es/~batanero/ARTICULOS/BataneroregularlectureICME.pdf&sa=X&scisig=AAGBfm3btVReL7r65V6GKEXgkSQKDJSdcw&oi=scholarr&ei=krx0T8maMsXb0QX_19TiDw&ved=0CCMQgAMoADAA (accessed 30th Mar 2012)

Batanero, C., Estepa, A., Godino, J. D. & Green, D. (1996). Intuitive strategies and preconceptions about association in contingency tables, *Journal for Research in Mathematics Education,* **27** (2), 151-169.

Batanero, C., Tauber, L. & Sánchez, L. (2004). Students' reasoning about the normal distribution. In J.B. Garfield & D. Ben-Zvi (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp.257-276). Dordrecht, The Netherlands: Kluwer.

Berg, B. L. (2004). *Qualitative Research Methods for the Social Sciences* (5th edition). Boston: Pearson Education, Inc.

Berg, C. & Phillips, D.G. (1994). An investigation of the relationship between logical thinking structures and the ability to construct and interpret line graphs. *Journal of Research in Science Teaching,* **31**, 323–344.

Bernard, T & Tall, D. (1997). Cognitive units, connections, and mathematical proof. In E. Pehkonen (Ed.), *Proceedings of the 21st Annual Conference for the Psychology of Mathematics Education* (**2**, pp.41-48). Jyväskylä, Finland: Gummerus.

Biggs, J. (1979). Individual differences in study processes and the quality of learning outcomes. *Higher Education,* **8**, 381-394. Amsterdam, The Netherlands: Elsevier Scientific Publishing Company.

Bluman, A.G. (2008). *Elementary Statistics: A Step by Step Approach* (4th edition). New York: McGraw-Hill.

Booth, L.R. (1981). Child methods in secondary mathematics. *Educational Studies in Mathematics*, **12**, 29-41.

Brase, G.L. (2009). Pictorial representation in statistical reasoning. *Applied Cognitive Psychology*, **23**, 369-381.

Breen, C. (1997). Exploring imagery in P, M and E. In E. Pehkonen (Ed.), *Proceedings of the 21st Annual Conference for the Psychology of Mathematics Education* (**2**, pp.97-104). Jyväskylä, Finland: Gummerus..

Bright, G.W. & Friel, S.N. (1996). Connecting stem plots to histograms. In E. Jakubowski, D. Watkins & H. Biske (Eds.), *Proceedings of the eighteenth annual meeting, North American Chapter of the International Group for the Psychology of Mathematics Education* (**1**, pp.235–240). Columbus, OH: ERIC Clearinghouse for

Science, Mathematics, and Environmental Education.

British Educational Research Association (2004). *Revised Ethical Guidelines for Educational Research*. Southwell, Notts: BERA.

Brown, T. & Liebling, H. (2005). *The Really Useful Maths Book: A Guide to Interactive Teaching*. Abingdon: Routledge.

Bryman, A.E. & Burgess R.G. (1999). *Qualitative Research* (4 volumes) (Eds.). London: Sage.

Carpenter, P.A. & Shah, P. (1998). A model of the perceptual and conceptual process in graph comprehension. *Journal of Experimental Psychology: Applied*, **4** (2), 75-100.

Chance, B., delMas, R. & Garfield, J. (2004). Reasoning about sampling distributions, In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp.295-324). Dordrecht, The Netherlands: Kluwer.

Charmaz, K. (2005). Grounded theory in the 21st Century: Application for advancing social justice studies, In N.K. Denzin & Y.S. Lincoln (2005) *Handbook of Qualitative Research* (3rd edition). Thousand Oaks: Sage.

Clements, M.A. (1984). Terence Tao. *Educational Studies in Mathematics*, **15**, 213-238.

Cleveland, W.S. (1985). *The Elements of Graphing Data*. Monterey, CA: Wadsworth.

Clough, P. & Nutbrown, C. (2007). *A Student's Guide to Methodology* (2nd edition). London: Sage.

Cohen, L., Manion, L. & Morrison, K. (2007). *Research Methods in Education* (6th edition). London: Routledge Falmer.

Corbin, J. & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3rd edition). Thousand Oaks: Sage.

Corter, J.E. & Zahner, D.C. (2007). Use of external visual representations in probability problem solving. *Statistics Education Research Journal*, **6** (1), 22-50, Online: http://www.stat.auckland.ac.nz/serj (accessed 30th March 2012).

Cortés, A. & Pfaff, N. (2000). Solving equations and inequations: operational invariants and methods constructed by students. *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education*, **2**, 193-200. Hiroshima, Japan: Nishiki Print Co., Ltd.

Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, **58**, 1–73.

Crowley, L. & Tall, D. (2006) Two students: why does one succeed and the other fail? In A. Simpson (ed.) *Retirement as Process and Concept: a Festschrift for Eddie Gray and David Tall* (pp.57-66). Karlova Univerzita v Praze, Pedagogicka Fakulta.

Curcio, F.R. (1981). *The Effect of Prior Knowledge, Reading and Mathematics Achievement, and Sex on Comprehending Mathematical Relationships Expressed in Graphs. Final Report*. Washington DC: National Institute of Education.

Curcio, F.R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, **18**, 382–393.

DelMas, R. C., Garfield, J. & Chance, B. L. (1999). A model of classroom research in action: developing simulation activities to improve students' statistical reasoning, *Journal of Statistics Education*, **7** (3) (n. p.). Online: http://www.amstat.org/publications/jse/secure/v7n3/delMas.cfm (accessed 30th March 2012).

Deutscher, I., Pestello, F.P. and Pestello, H.F.G. (1993). *Sentiments and Acts*. New York: Walter de Gruyter, Inc.

Diezman, C.M. (2005). Primary students' knowledge of the properties of spatial-oriented diagrams. In H.L. Chick & J.L. Vincent (Eds.), *Proceedings of the 29th PME International Conference*, **2**, 281-288.

Dreyfus, T. (1991). On the status of visual reasoning in mathematics and mathematics education. In F. Furinghetti (Ed.), *Proceedings of the 15th PME International Conference*, **1**, 33-48. Assisi, Italy: Program Committee of the 15th PME

Conference.

Dreyfus, T. (2006). Linking theories in mathematics education. In A. Simpson (Ed.) *Retirement as Process and Concept: A Festschrift for Eddie Gray and David Tall* (pp.67-76). Karlova Univerzita v Praze, Pedagogicka Fakulta.

Dunham, P.H. & Osborne, A. (1991). Learning how to see: Students' graphing difficulties. *Focus on Learning Problems in Mathematics*, **13** (4), 35–49.

Duval, R. (1999). Representation, vision and visualisation: Cognitive functions in mathematical thinking. Basic issues for learning. In F. Hitt & M. Santos (Eds.), *Proceedings of the 21st North American PME Conference*, **1**, 3-26.

Earley, M.A. (2001). *Improving statistics education through simulations: The case of the sampling distribution*. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association. Chicago, Illinois, October 24-27.

Eisenberg, T. & Dreyfus, T. (1991). On the reluctance to visualise in mathematics. In W. Zimmerman & S. Cunningham (Eds.), *Visualization in Teaching and Learning Mathematics* (pp.25-37). Washington, DC, USA: Mathematical Association of America.

Elbrink, M. (2008). Analyzing and addressing common mathematical errors in secondary education. *B.S. Undergraduate Mathematics Exchange*, **5** (1), 2-4.

Falk, R. (1988). Conditional probabilities: Insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the second international conference on the teaching of statistics* (ICOTS-2) (pp.292-297). Victoria BC, Canada: University of Victoria.

Falk, R. & Greenbaum, C.W. (1995). Significance test die hard: The amazing persistence of a probabilistic misconception. *Theory and Psychology*, **5**, 75-98.

Fey, J.T. (1990). Quantity. In L. A. Steen (Ed.), *On the Shoulders of Giants* (pp.61–99). Washington, DC: National Academy Press.

Foster, P. (1996). *Observing Schools: A Methodological Guide*. London: Paul Chapman Publishing Ltd.

Freudenthal, H. (1983). *Didactical Phenomenology of Mathematical Structures*. Dordrecht, Holland: D. Reidal Publishing Company.

Friel, S.N, Curcio, F.R. & Bright, G.W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, **32**, 124-158.

Gal, I. (1993). Reaching out: Some issues and dilemmas in expanding statistics education. In L. Pereira-Mendoza (Ed.), *Introducing data-analysis in the schools: Who should teach it and how?* (pp.189–203). Voorburg, The Netherlands: International Statistics Institute.

Gal, I. (2002). Adults' statistical literacy: Meaning, components, responsibilities. *The International Statistics Review*, **70** (1), 1-25.

Garfield, J.B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, **2** (1), 22-38.

Garfield, J.B. & Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning*: *Connecting Research and Teaching Practice*. Springer Science + Business Media B.V.

Giaquinto, M. (1994). Epistemology of visual thinking in elementary real analysis. *The British Journal for the Philosophy of Science*, **45** (3), 789-813.

Giaquinto, M. (2005). Mathematical activity. In P. Mancosu, K.F. Jorgensen & S.A. Pedersen (Eds.), *Visualization, Explanation and Reasoning Styles in Mathematics* (pp.75-87). Dordrecht, The Netherlands: Springer.

Ginsburg, H. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the Learning of Mathematics*, **1** (3), 4-11.

Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice.* New York: Cambridge University Press.

Godino, J.D. & Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. In A. Sierpinska & J. Kilpatrick

(Eds.), *Mathematics Education as a Research Domain: A Search for Identity* (pp.177-195). Dordrecht, The Netherlands: Kluwer Academic.

González-Martín, A. S. & Camacho, M. (2004). Legitimization of the graphic register in problem solving at the undergraduate level. The case of the improper integral. *Proceedings of the 28th Conference of the PME*, **2**, 479-486.

Graham, A. (1987). *Statistical Investigations in the Secondary School*. New York: Cambridge University Press.

Gray, E. (1999). Spatial strategies and visualisation. In O. Zaslavsky (Ed.) *Proceedings of the 23rd PME International Conference*, **1**, 235-242. Haifa, Israel.

Gray, E. & Doritou, M. (2006) Estimating on a number line: An alternative view. In A. Simpson (ed.) *Retirement as Process and Concept: A Festschrift for Eddie Gray and David Tall* (pp.67-75). Karlova Univerzita v Praze, Pedagogicka Fakulta.

Guzmán, M. de (2002). The role of visualization in the teaching and learning of mathematical analysis. *Proceedings of the 2nd International Conference on the Teaching of Mathematics (at the undergraduate level)* (n. p.). Crete, Greece: University of Crete.

Harry, B., Sturges, K.M., & Klingner, J.K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational Researcher,* **34** (2), 3-13.

Healy, L. & Hoyles, C. (1996). Seeing, doing and expressing: An evaluation of task sequences for supporting algebraic thinking. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the 20th PME International Conference*, **3**, 67-74. Valencia, Spain: University of Valencia.

Hilbert, J., Carpenter, T.P., Fennema, E., Fuson, K.C., Wearne, D., Murray, H. Oliver, A. & Human, P. (1997). *Making Sense: Teaching and Learning Mathematics with Understanding*. Portsmouth, NH: Heinemann.

Hoffman, D. (1987). The computer-aided discovery of new embedded minimal surfaces, *The Mathematical Intelligencer*, **9**, 8-21.

Jolliffe, F.R. (1991). Assessment of the understanding of statistical concepts. In D. Vere-Jones (Ed.), *Proceedings of the third international conference on teaching statistics* (Vol. 1, pp.461–466). Voorburg, The Netherlands: International Statistical Institute.

Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge Press.

Khera, N., Stroobant,T., Primhak, R.A., Gupta, R. & Davies, H. (2001). Training the ideal hospital doctor: the specialist registrars' perspective. *Medical Education*, 35(10), 957–966.

Kirsch, I.S., Jungeblut, S.S. & Mosenthal, P.M. (1998). The measurement of adult literacy. In S.T. Murray, I.S. Kirsch & L.B Jenkins (Eds.), *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey* (pp.105-134). Washington DC: National Center for Education Statistics, U.S. Department of Education.

Konold, C. (1989). An outbreak of belief in independence? In the *eleventh annual meeting of the North American chapter of the international group for the psychology of mathematics education* (Vol. 2, pp.203-209). New Brunswick, NJ: Center for mathematics, Science and Computer Education.

Konold, C. (1991). Understanding students' beliefs about probability. In E.V. Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp.139-156). Amsterdam, The Netherlands: Kluwer Academic Publishers.

Kosslyn, S.M. (1985). Graphics and human information processing: A review of five books. *Journal of the American Statistical Association,* **80**, 499–512.

Kosslyn, S.M. (1994). *Elements of Graph Design*. New York: W. H. Freeman.

Kuo, S & Shi, S. (2006). *Statistics* (6th edition Chinese textbook). Taipei, Taiwan: Jolihi.

Larkin, J.H. & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, **11**, 65-99.

Lecoutre, B. (1998). Teaching Bayesian methods for experimental data analysis. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee & W.K.Wong (Eds.) *Proceedings of the fifth international Conference on Teaching of Statistics*, **1**, 239-244. Singapore: International Statistical Institute.

Legge, G.E., Gu, Y. & Luebker, A. (1989). Efficiency of graphical perception. *Perception and Psychophysics,* **46,** 365–374.

Legutko, M. (2008). An analysis of students' mathematical errors in the teaching-research process. In B. Czarnocha (Ed.) *Handbook of Mathematics Teaching Research: Teaching Experience – A Tool for Teacher-Researchers* (pp.141-152). Poland: University of Rzeszów.

Lima, R.N. & Tall, D.O. (2006a). The concept of equations: What have students met before? *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education*, **4**, 233-240. Prague, Czech Republic: Atelier Guimaec, s.r.o.

Lima, R.N. & Tall, D.O. (2006b). What does equation mean? A brainstorm of the concept. *Proceedings of the 3rd International Conference on the Teaching of Mathematics*, 1-5. Istanbul, Turkey.

Lindberg, E., Rosenqvist, U. (2003). *The staff's views on what's the job – the starting point for quality improvement in health care.* International journal of health care quality assurance incorporating leadership in health services, vol. 16, ss. 47-54.

Liu, Y. (2005). *Teachers' Understanding of Probability and Statistical Inference and their Implications for Professional Development*. Ph.D. Dissertation, Online: http://www.stat.auckland.ac.nz/~iase/publications/dissertations/05.liu.Dissertation.pdf (accessed 30th March 2012).

Liu, M., Chen, C., Shi, H., Wang, K., Liu, K., Lin, Z., Wu, y. & Lin, H. (2005). *Statistics* (Chinese textbook). Taipei, Taiwan: Chwa.

Lovett, M.C. (2001). A collaborative convergence on studying reasoning process: A case study in statistics. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp.347-384). Mahwah, NJ: Erlbaum.

Lowrie, T. (2001). The influence of visual representations on mathematical problem solving and numeracy performance. *24th Annual MERGA Conference,* Sydney.

Madriz, E. (1998). Using focus groups with lower socioeconomic status Latina woman. *Qualitative Inquiry*, **24** (1), 114-128.

Madriz, E. (2003). Focus groups in feminist research. In N.K. Denzin & Y.S. Lincoln (Eds.), *Collecting and Interpreting Qualitative Materials* (2nd edition) (pp.363-388). Thousand Oaks, CA: Sage.

Mancosu, P. (2005). Visualization in logic and mathematics. In P. Mancosu, K.F. Jorgensen & S. A. Pedersen (Eds.), *Visualization, Explanation and Reasoning Styles in Mathematics* (pp.13-30). Dordrecht, Netherlands: Springer.

Maxwell, E.A. (1959). *Fallacies in Mathematics*. New York: Cambridge University Press.

Mellin-Olsen, S. (1987). *The Politics of Mathematics Education*. Dordrecht, The Netherlands: D. Reidal Publishing Company.

Mertens, D.M. (1998). *Research Methods in Education and Psychology: Integrating Diversity with Quantitative & Qualitative Approaches*. London: Sage.

Miller, B. & Bell, L. (2002). Consenting to what? Issues of access, gate-keeping and ‘informed’ consent. In M. Mauthner, M. Birch, J. Jessop, & T. Miller, *Ethics in Qualitative Research* (pp.53-69). London: Sage.

Moore, D.S. (1991). *Statistics: Concepts and controversies* (3rd edition). New York: W. H. Freeman.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

Pratt, D. & Ainley, J. (2008). Introducing the special issue on informal inferential reasoning. *Statistics Education Research Journal*, **7** (2), 3-4. International Association for Statistical Education (IASE/ISI)

Pfannkuch, M. (2006). Informal inferential reasoning. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil (pp.1-6). Voorburg, The Netherlands: International Statistical Institute.

Presmeg, N.C. (2006). Research on visualization in learning and teaching mathematics. In A. Gutiérrez & P. Boero (Eds.), *Handbook of Research on the Psychology of Mathematics Education: Past, Present and Future* (pp.205-235). Rotterdam, The Netherlands: Sense Publisher.

Presmeg, N. C. & Bergsten, C. (1995). Preference for visual methods: An international study. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19th PME International Conference*, **3**, 58-65. Recife, Brazil: Program committee of the 19th PME conference.

Pugalee, D.K. (2004). A comparison of verbal and written descriptions of students' problem solving process. *Educational Studies in Mathematics*, **55**, 27-47.

Quilici, J.L. & Mayer, R.E. (2002). Teaching students to recognize structural similarities between statistics word problems. *Applied Cognitive Psychology*, **16** (3), 325-342.

Rangecroft, M. (1994). Graphwork—Developing a progression. In D. Green (Ed.), *Teaching Statistics at its Best* (pp.7–12). Sheffield, England: The Teaching Statistics Trust.

Reid, A., and Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education*, **10** (2). Online: http://www.amstat.org/publications/jse/v10n2/reid.html (accessed 30th March 2012).

Rossman, A.J. & Chance, B.L. (1999). Teaching the Reasoning of Statistical Inference: A "Top Ten" List. *College Mathematics Journal*, **30** (4), 297-305. Online:

http://www.rossmanchance.com/papers/topten.html (accessed 30th March 2012).

Roth, W.-M. & McGinn, M.K. (1997). Graphing: Cognitive ability or practice? *Science Education,* **81**, 91–106.

Schau, C. & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp.91-104). Amsterdam, The Netherlands: IOS Press.

Schwartz, D. L. & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, **22** (2), 129-184

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications*. Hillsdale, NJ: Erlbaum.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching learning* (pp.465-494). New York: Macmillan.

Shaughnessy, J.M. (2007). Research on statistics learning and reasoning. In F.K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp.957-1009). Charlotte: Information Age Publishing Inc.

Siegler, R.S. & Booth, J.L. (2004). Development of numerical estimation in young children. *Children Development*, **75**, 428-444.

Sierpinska, A. (1994). *Understanding in Mathematics*. London: the Falmer Press.

Simkin, D. & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association*, **82**, 454–465.

Sloman, S. A., Over, D., Slovak, L. & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, **91**, 296–309.

Stix, A. (1994). Pic-Jour math: Pictorial journal writing in mathematics. *Arithmetic Teacher*, **41** (5), 264-269

Sutherland, R. (2007). *Teaching for Learning Mathematics*. Maidenhead: Open University Press.

Tall, D.O. (2002). Continuities and discontinuities in long-term learning schemas (Reflecting on how relational understanding may be instrumental in creating learning problems). In D.Tall & M. Thomas (Eds.), *Intelligence, Learning and Understanding in Mathematics – A tribute to Richard Skemp* (pp.151-177). Flaxton, Australia: Post Pressed.

Tall, D.O. (2004). Thinking through three worlds of mathematics, *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, Bergen, Norway, **4**, 281–288. International Group for the Psychology of Mathematics Education.

Tall, D. & Vinner, S. (1981). Concept images and concept definitions in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, **12**, 151-169.

Tarr, J.E. & Shaughnessy, J.M. (2007). Student performance in data analysis, statistics and probability. In P. Kloosterman & F. Lester (Eds.), *Results and Interpretations of the 2003 Mathematics assessment* (pp.139-168). Reston, VA: NCTM.

Thompson, P.W., Cheepurupalli, R., Hardin, B., Lienert, C. & Selden, A. (2010). Cultivating symbol sense in your calculus class. IM&E workshop March 27-29, 1-5.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Verschaffel, L., Greer, B. & Torbeyns, J. (2006). Numerical thinking. In A. Gutiérrez & P. Boero (Eds.), *Handbook of Research on the Psychology of Mathematics Education: Past, Present and Future* (pp.51-81). Rotterdam, The Netherlands: Sense Publisher.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences*, **22**, 219–240.

Vlassis, J. (2002). The balance model: Hindrance or support for the solving of linear equations with one unknown. *Educational Studies in Mathematics*, **49**, 341-359. Kluwer Academic Publishers. Dordrecht, the Netherlands: Springer.

Wainer, H. (1992). Understanding Graphs and Tables. *Educational Researcher,* **21** (1), 14–23.

Watson, J.M. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp.277-294) Dordrecht, The Netherlands: Kluwer.

Watson, J.M. & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, **2**, 3-46.

Wheatley, G.H. & Brown, D. (1994). The construction and representation of images in mathematical activity. In J.P. Ponte & J.F. Matos (Eds.), *Proceedings of the 18th PME International Conference*, **1**, 81. Lisbon, Portugal: Program committee of the 18th PME conference

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal,* **5** (2), 10-26.

Wild, C.J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry, *International Statistical Review*, **67**, 221-248.

Wisenbaker, J. & Scott, J. (1997). *Modeling aspects of students' attitudes and achievement in introductory statistics courses*. Paper presented at annual meeting of the American Educational Research Association (AERA), Chicago, IL.

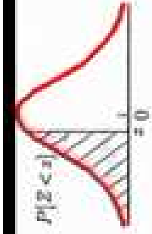
Wonnacott, T.H. & Wonnacott, R.J. (1990). *Introductory Statistics* (5th edition). New York: John Wiley & Sons.

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K. & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college Level? A review of the literature. *Journal of Statistics Education* **16** (2), 1-23.

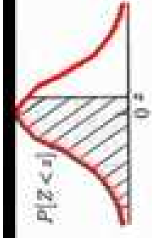
Zieffler, A, Garfield, J., Delmas, R. & Reading, C. A framework to support reasoning on informal inferential reasoning. *Statistics Education Research Journal*, **7** (2), 40-5.

APPENDICES

1.1. The table of normal distribution used in my statistics course.



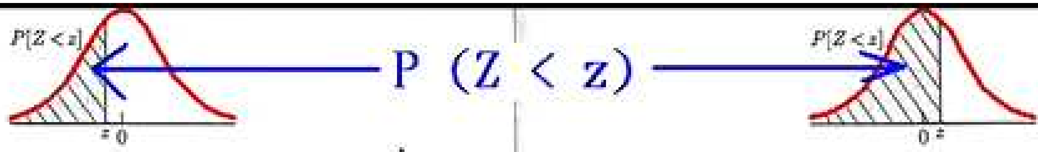
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

此表採錄於 Johnson, R.A. and Bhattacharyya, G.K. (1996), *STATISTICS*, 3 edition, John Wiley and Sons, Inc.

1.2. Symbolic explanation of the shadow area.

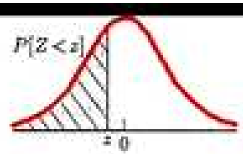


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

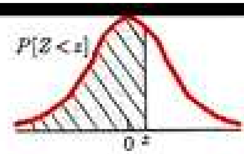
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9845	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

此表摘錄於 Johnson, R.A. and Bhattacharyya, G.K. (1996). *STATISTICS*, 3 edition, John Wiley and Sons, Inc.

1.3. Example of the locations of z score and probability in the table.



$P(Z < -3.50) = 0.0002$

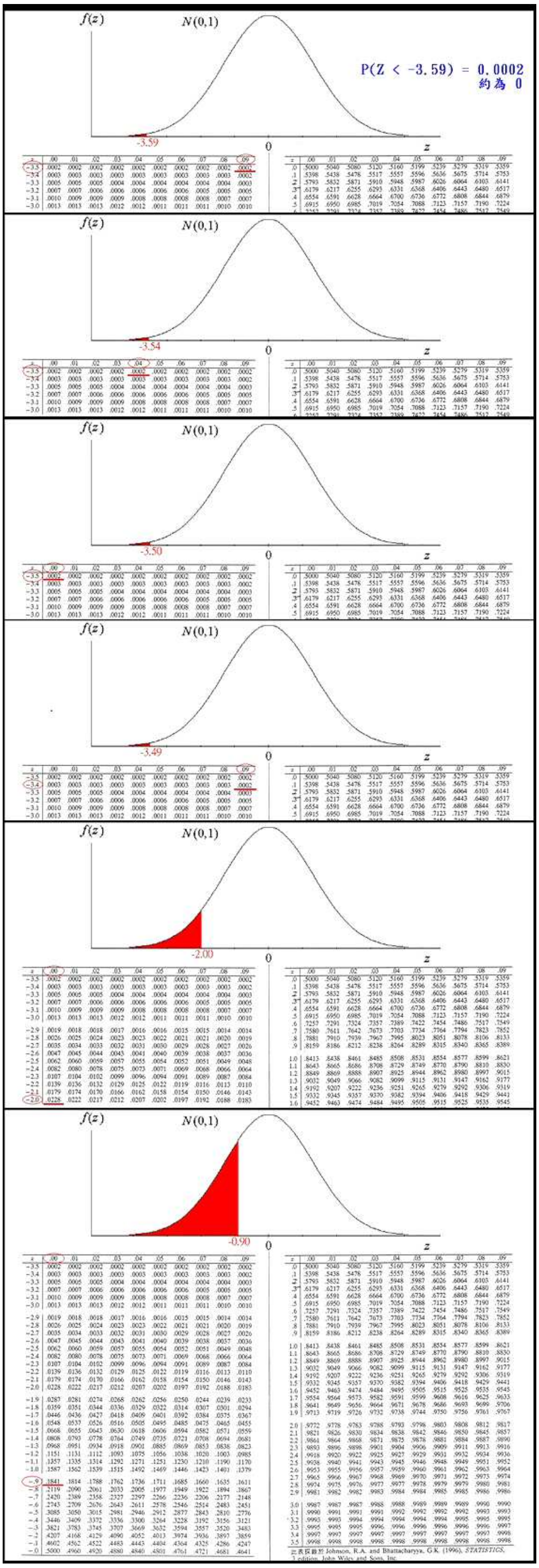


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

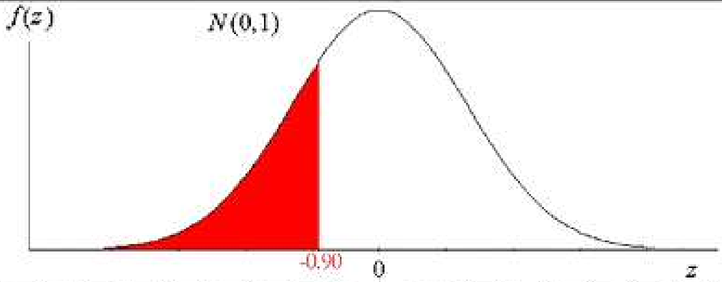
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9845	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

此表採錄於 Johnson, R.A. and Bhattacharyya, G.K. (1996), *STATISTICS*, 3 edition, John Wiley and Sons, Inc.

1.4. Increasing z scores and probabilities in the table and expanding areas.



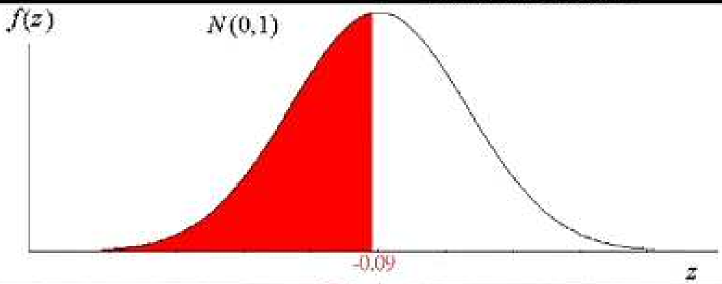
1.5. Comparison of confusing values.



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

此表摘錄於 Johnson, R.A. and Bhattacharyya, G.K. (1996), *STATISTICS*, 3 edition, John Wiley and Sons, Inc.

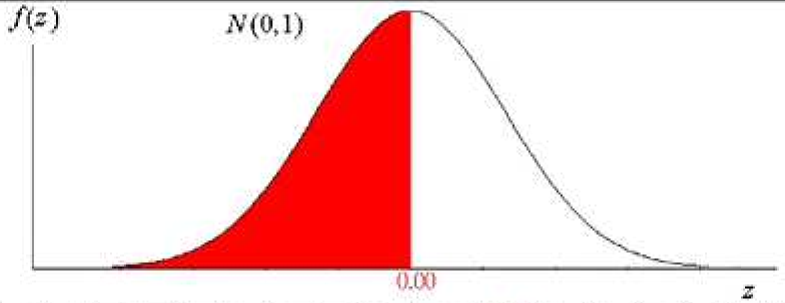


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

此表摘錄於 Johnson, R.A. and Bhattacharyya, G.K. (1996), *STATISTICS*, 3 edition, John Wiley and Sons, Inc.

1.6. Link P(Z < -0.00) and P(Z < 0.00).

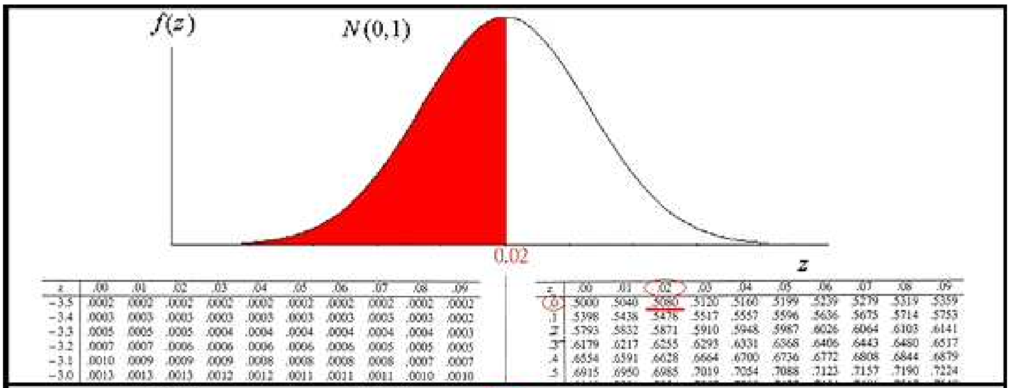


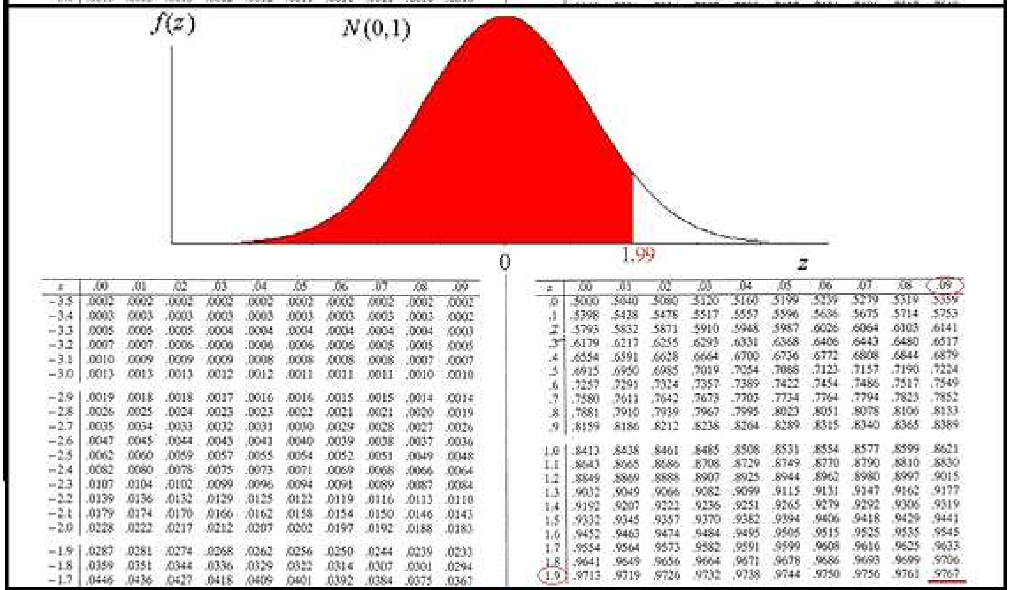
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

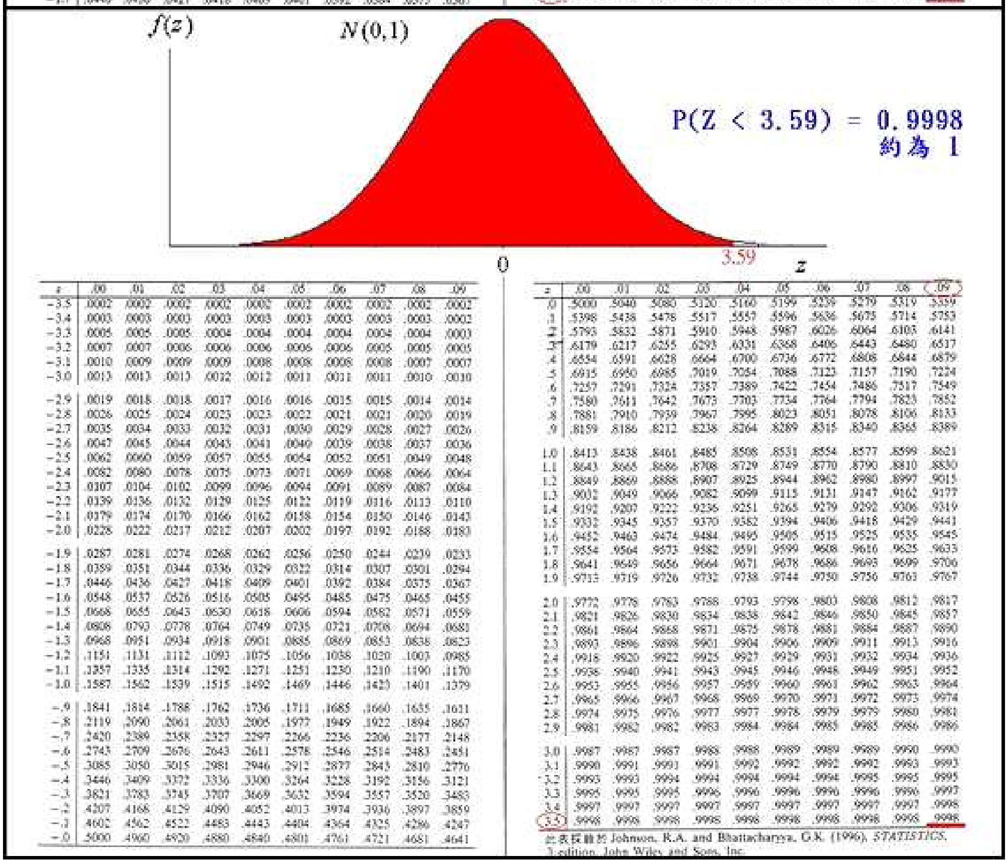
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

此表採錄於 Johnson, R.A. and Bhattacharyya, G.K. (1996), *STATISTICS*, 3 edition, John Wiley and Sons, Inc.

1.7. Increasing z scores and probabilities in the table and expanding areas.

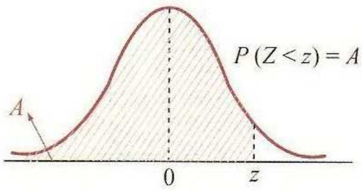






2.1. Table A of normal distribution (negative z scores and positive z scores).

■ 表 A–4 標準常態分配



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
− 3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
− 3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
− 3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
− 3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
− 3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
− 3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
− 3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
− 3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
− 3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
− 3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
− 2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
− 2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
− 2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
− 2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
− 2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
− 2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
− 2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
− 2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
− 2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
− 2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
− 1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
− 1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0370	.0301	.0294
− 1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
− 1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
− 1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
− 1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
− 1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
− 1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
− 1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
− 1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
− .9	.1841	.1841	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
− .8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1992	.1894	.1867
− .7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
− .6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
− .5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
− .4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
− .3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
− .2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
− .1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
− .0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

■ 表 A-4（續一） 標準常態分配

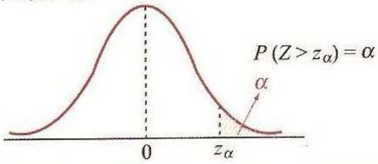
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.9708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.8	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

572 Statistics

(Kuo and Shi, 2006, pp.571-572).

2.2. Table B of normal distribution (positive z scores).

■ 表 A–4（續二） 標準常態分配

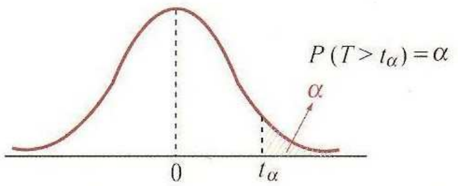


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1304	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

(Kuo and Shi, 2006, p.573).

2.3. Table of t distribution.

■ 表 A–5 *t* 分配



v	α = .25	.2	.1	.05	.025	.01	.00833
1	1.000	1.376	3.078	6.314	12.706	31.821	38.190
2	.816	1.061	1.886	2.920	4.303	6.965	7.469
3	.765	.978	1.638	2.353	3.182	4.541	4.857
4	.741	.941	1.533	2.132	2.776	3.747	3.961
5	.727	.920	1.476	2.015	2.571	3.365	3.534
6	.718	.906	1.440	1.943	2.447	3.143	3.287
7	.711	.897	1.415	1.895	2.365	2.998	3.128
8	.706	.889	1.397	1.860	2.306	2.896	3.016
9	.700	.883	1.383	1.833	2.262	2.821	2.933
10	.700	.879	1.372	1.812	2.228	2.764	2.870
11	.697	.876	1.363	1.796	2.201	2.718	2.820
12	.695	.873	1.356	1.782	2.179	2.681	2.779
13	.694	.870	1.350	1.771	2.160	2.650	2.746
14	.692	.868	1.345	1.761	2.145	2.624	2.718
15	.691	.866	1.341	1.753	2.131	2.602	2.694
16	.690	.865	1.337	1.746	2.120	2.583	2.673
17	.689	.863	1.333	1.740	2.110	2.567	2.655
18	.688	.862	1.330	1.734	2.101	2.552	2.639
19	.688	.861	1.328	1.729	2.093	2.539	2.625
20	.687	.860	1.325	1.725	2.086	2.528	2.613
21	.686	.859	1.323	1.721	2.080	2.518	2.601
22	.686	.858	1.321	1.717	2.074	2.508	2.591
23	.685	.858	1.319	1.714	2.069	2.500	2.582
24	.685	.857	1.318	1.711	2.064	2.492	2.574
25	.684	.856	1.316	1.708	2.060	2.485	2.566
26	.684	.856	1.315	1.706	2.056	2.479	2.559
27	.684	.855	1.314	1.703	2.050	2.473	2.552
28	.683	.854	1.313	1.701	2.048	2.467	2.546
29	.683	.854	1.311	1.699	2.045	2.462	2.546
30	.683	.854	1.310	1.697	2.042	2.457	2.536
40	.681	.851	1.303	1.684	2.021	2.423	2.499
60	.679	.848	1.296	1.671	2.000	2.390	2.463
120	.677	.844	1.289	1.658	1.980	2.358	2.428
∞	.674	.842	1.282	1.645	1.960	2.326	2.394

■ 表 A–5（續） t 分配

v	.00625	.005	.0025	.001	.0005	.00025	.0001	.00005
1	50.923	63.657	127.21	318.309	363.619	1,273.239	3,183.099	6,366.198
2	8.860	9.925	14.089	22.327	31.598	44.705	70.700	99.992
3	5.392	5.841	7.453	10.214	12.924	16.326	22.204	28.000
4	4.315	4.604	5.598	7.173	8.610	10.306	13.034	15.544
5	3.810	4.032	4.773	5.893	6.869	7.976	9.678	11.178
6	3.521	3.707	4.317	5.208	5.959	6.788	8.025	9.082
7	3.335	3.499	4.029	4.785	5.408	6.082	7.063	7.885
8	3.206	3.355	3.833	4.501	5.041	5.618	6.442	7.120
9	3.111	3.250	3.690	4.297	4.781	5.291	6.010	6.594
10	3.038	3.169	3.581	4.144	4.587	5.049	5.694	6.211
11	2.981	3.106	3.497	4.025	4.437	4.863	5.453	5.921
12	2.934	3.055	3.428	3.930	4.318	4.716	5.263	5.694
13	2.896	3.012	3.372	3.852	4.221	4.597	5.111	5.513
14	2.864	2.977	3.326	3.787	4.140	4.499	4.985	5.363
15	2.837	2.947	3.286	3.733	4.073	4.417	4.880	5.239
16	2.813	2.921	3.252	3.686	4.015	4.346	4.791	5.134
17	2.793	2.898	3.223	3.646	3.965	4.286	4.714	5.044
18	2.775	2.878	3.197	3.610	3.922	4.233	4.648	4.966
19	2.759	2.861	3.174	3.579	3.883	4.187	4.590	4.897
20	2.774	2.845	3.153	3.552	3.850	4.146	4.539	4.837
21	2.732	2.831	3.135	3.527	3.819	4.110	4.493	4.784
22	2.720	2.819	3.119	3.505	3.792	4.077	4.452	4.736
23	2.710	2.807	3.104	3.485	3.768	4.048	4.415	4.693
24	2.700	2.797	3.090	3.467	3.745	4.021	4.382	4.654
25	2.692	2.787	3.078	3.450	3.725	3.997	4.352	4.619
26	2.684	2.779	3.067	3.435	3.707	3.974	4.324	4.587
27	2.676	2.771	3.057	3.421	3.690	3.954	4.299	4.558
28	2.669	2.763	3.047	3.408	3.674	3.935	4.275	4.530
29	2.663	2.756	3.038	3.396	3.659	3.918	4.254	4.506
30	2.657	2.750	3.030	3.385	3.646	3.902	4.234	4.482
40	2.616	2.704	2.971	3.307	3.551	3.788	4.094	4.321
60	2.575	2.660	2.915	3.232	3.460	3.681	3.962	4.169
120	2.536	2.617	2.871	3.174	3.390	3.598	3.862	4.053
∞	2.498	2.576	2.807	3.090	3.291	3.481	3.719	3.891

Statistics 575

(Kuo and Shi, 2006, pp.574-575).

4.1 Blank consent form and information sheet for the teacher.

Dear Sir

I am a PhD student in the School of Education and Lifelong Learning at the University of East Anglia (Norwich, UK) and I am currently collecting data for my doctoral thesis. One of my research areas is how visualisation helps statistics teaching in statistical inference, so I would like to invite you to participate in my research during the 8^{th} week to the 18^{th} week of this semester.

In many statistics courses, mathematics and formulas take a large part of the content, and I noticed that many students had difficulties when they learned this subject. In order to help student learn statistics better, I will focus on the influence of visualization in teaching and learning.

During the next three months, I would like to observe, record, and make notes in your class, collect your class material, and collect student participants' examination papers from you. I would also like to invite you to interviews which will be audio-recorded.

These collected data will not be used to judge your teaching. They will be totally confidential, and will not be released to anyone else, except for you and me. Your name, title and other clues which can be used to trace back to you will not be used in my paper; alternatively I will use pseudonyms to represent different participants.

The interviews will be arranged during this period, and you will be interviewed for four times every three weeks (initially arranged in week 9, 12, 15 and 18 of the second semester). Each interview will last between thirty to forty minutes and you will be interviewed personally.

Participating in this research is totally voluntary, and you can also withdraw at any time. If you have any queries before or after participating in my research, you are always welcome to contact me. This data collection is entirely and only for the purpose of my academic research and PhD thesis.

Thank you

Che-yu Shen (EDU PhD student)
E-mail: c.shen@uea.ac.uk
Mobile: 093xxxx357

Participant's Consent

I am willing to participate in the research of Mr Che-yu Shen. I will allow Mr Shen to observe, record, and make notes in my class, collect my class material, and student participants' examination papers in statistics course, and be interviewed for his research, under the UEA ethical principles. I also understand my right to participate in and withdraw from this research.

Signature____________________________ Date___________________________

4.2 Blank consent form and information sheet for the students.

Dear student

I am a PhD student in the School of Education and Lifelong Learning at the University of East Anglia (Norwich, UK) and I am currently collecting data for my doctoral thesis. One of my research areas is how visualisation helps statistics learning in statistical inference, so I would like to invite you to participate in my research during the 8th week to the 18th week of this semester.

In many statistics courses, mathematics and formulas take a large part of the content, and I noticed that many students had difficulties when they learned this subject. In order to help student learn statistics better, I will focus on the influence of visualization in learning.

During the next three months, I would like to collect your notes, homework and examination papers. I would also like to invite you to interviews which will be audio-recorded.

These collected data will neither be used to judge your performance in class, nor influence your score in the end of semester. They will be totally confidential, and will not be released to anyone else, except for each of you and me. Your name, class and year will not be used in my paper; alternatively I will use pseudonyms to represent different students.

The interviews will be arranged during this period, and you will be interviewed for three times every three weeks (initially arranged in week 10, 13, and 16, or 11, 14 and 17 of the second semester). Each interview will last between thirty to forty minutes and you will be interviewed personally, but you can also form a cooperative learning group and be interviewed as a group.

Participating in this research is totally voluntary, and participants can also withdraw at any time. If you have any queries before or after participating in my research, you are always welcome to contact me. This data collection is entirely and only for the purpose of my academic research and PhD thesis.

Thank you

Che-yu Shen (EDU PhD student)
E-mail: c.shen@uea.ac.uk
Mobile: 093xxxx357

Participant's Consent

I am willing to participate in the research of Mr Che-yu Shen. I will allow Mr Shen to collect my notes, homework and examination papers in statistics course and be interviewed for his research, under the UEA ethical principles. I also understand my right to participate in and withdraw from this research.

Signature____________________________ Date___________________________

4.3 Ethics form

School of Education & Lifelong Learning Research Ethics Committee

Guidance for Students and their Supervisors

This document is intended to provide outline guidance for students who are undertaking a piece of research involving human participants, as a part of their programme of study. It sets out the ways in which the school is implementing the Research Ethics Policy of the University, and the procedures that students and supervisors should follow in obtaining ethics approval for the planned research. It should be noted that the University Policy was revised on 21st June 2006, and as a result of this the School has been required to amend its own procedures. Thus, this document supercedes any existing forms and protocols in use. The guidance document should be read in conjunction with the “University Research Ethics Policy, Principle and Procedures” available at:

http://www.uea.ac.uk/rbs/rso/research_ethics/index.htm

In particular, it should be noted that:

> All University members of staff and University-registered students (i.e. postgraduate research, postgraduate taught and undergraduate students) who plan to undertake research that falls under the scope of the Ethics Principles in the Policy must obtain ethics approval for the planned research prior to the involvement of the participants via the appropriate ethics review procedure. The Procedures also apply to all individuals who are performing research which is funded or managed by the University, be this on or off University premises.

And:

> Research involving human participants ("participants") is defined broadly to include research that:
>
> • directly involves people in the research activities, through their physical participation. This may be invasive (e.g. surgery) or non-invasive research (e.g. interviews, questionnaires, surveys, observational research) and may require the active or passive involvement of a person;

- indirectly involves people in the research activities, through their provision of or access to personal data and/or tissue;
- Involves people on behalf of others (e.g. legal guardians of children and the psychologically or physically impaired and supervisors of people under controlled environments (e.g. prisoners, school pupils)).

Thus any research involving human participants requires ethical review and approval by the appropriate committee. The starting point for this will usually be the EDU Ethics Committee, although certain proposals may also require approval by other committees inside or outside of UEA. Also, although this document speaks to the procedural requirements of the University, it is important that students also consider the broader ethical implications of their work, and that they continue to review these with their supervisor as the research progresses.

1. Research that will result in a dissertation or thesis:

Students are required to prepare a research proposal including a discussion of the ethical issues involved in the proposed study, together with a Participant Information Sheet and an Informed Consent Form.

For students registered for a research degree (e.g. Ph.D., Ed.D., M.Phil., M.Res.) the proposal should be approximately two to three thousand words in length.

For students registered for a taught degree (e.g. M.A.) the proposal should be approximately three to four hundred words in length.

Students must discuss their proposal with their supervisor and then submit, by email only, one copy of this proposal. This must be accompanied by the form – "Application for Ethical Approval of a Research Project" (attached below, and also available on the EDU Research website and Blackboard).

The Chair of the Committee will then determine the procedure through which ethical approval will be granted. In many cases, where the project is determined to be of minimal risk, the proposal will not be seen by the full committee, but will be approved by the Chair or Deputy Chair in an expedited manner.

2. Research that is a taught student project (e.g. all course assignments):

Assignments that have a research element, but will not result in a dissertation or thesis do not automatically require ethical approval from the Committee, though students are still required to consider the ethical issues arising from their work and to adhere to the University Policy in terms of their obligations to participants. Students should discuss the ethical issues of their proposed project with the course tutor to ensure that the proposed work does not raise ethical issues that cannot be adequately dealt with through the process of obtaining informed consent from participants. If the project does raise more serious ethical concerns then the course tutor and/or student should consult with the Chair of the Ethics Committee, and if necessary apply for ethical approval from the Committee.

UNIVERSITY OF EAST ANGLIA
FACULTY OF SOCIAL SCIENCES
School of Education & Lifelong Learning

RESEARCH ETHICS COMMITTEE

APPLICATION FOR ETHICAL APPROVAL OF A RESEARCH PROJECT

This form is for all staff and students in the School of Education who are planning research that requires ethical approval. Applicants are advised to consult the school and university guidelines before preparing their application.
Completed applications (including the required attachments) must be submitted electronically to Dawn Corby d.corby@uea.ac.uk

1. APPLICANT DETAILS	
Name:	CHE-YU SHEN
School:	EDUCATION AND LIFELONG LEARNING
Current Status (delete as applicable):	**PGR Student**
If Student, name of primary supervisor and programme of study:	
	PAOLA IANNONE
Contact telephone number:	01603591007
Email address:	P.IANNONE@UEA.AC.UK

2. PROPOSED RESEARCH PROJECT DETAILS:	
Title:	HOW CAN VISUALIZATION HELP STATISTICS TEACHING & LEARNING IN STATISTICAL INFERENCE?
Start/End Dates:	APR 09 – SEP 10

3. FUNDER DETAILS (IF APPLICABLE):	
Funder:	
	Has funding been applied for? YES NO Application Date:
	Has funding been awarded? YES NO
Will ethical approval also be sought for this project from another source? YES NO	
	If "yes" what is this source?

3. DECLARATION:

I am satisfied that all ethical issues have been identified and that satisfactory procedures are in place to deal with those issues in this research project. I will abide by the procedures described in this form.

Name of Applicant:	CHE-YU SHEN
Date:	2ND APR 09

Supervisor declaration (for student research only)

I have discussed the ethics of the proposed research with the student and am satisfied that all ethical issues have been identified and that satisfactory procedures are in place to deal with those issues in this research project.

Name of Supervisor:	
Date:	

4. ATTACHMENTS:

The following should be attached to your application as necessary – please indicate if attached and list any additional materials:

▣ Research Proposal ▣ Project Information Sheet (for participants)
▣ Participant Consent Form(s) ☐ Other Supporting Documents

FOR ADMINISTRATIVE USE ONLY

Considered by Chair: (Date)
Considered at Committee Meeting: (Date)
Minute reference:

Recommendation:

Accept		**Amend and Resubmit**	
Amend		**Reject**	

Comments:

4.4 Ethical approval

The study was ethical approved by the chair of committee via e-mail.

Ethical Approval - Shen

Lyndon Martin [Lyndon.Martin@uea.ac.uk]
Sent: Wednesday, April 15, 2009 12:49 PM
To: Shen Che Mr (EDU)
Cc: Iannone Paola Dr (EDU)

Dear Leo

Thank you for your application for ethical review. As Chair of the Committee I am able to approve your study as proposed. I hope the research goes well.

One minor point, both consent forms are titled 'student' whereas one should be 'teacher'. There is no need to send me the amended versions.

Best

Lyndon

On 4/10/09 4:28 PM, "C.Shen@uea.ac.uk" <C.Shen@uea.ac.uk> wrote:

> hello Lyndon
>
> I am very sorry to send you teacher and students' consent forms so late
> I am in China now and I have could not link to UEA websites in the several
> days
>
> I put the ethics form signed by Paola and me in your pigeon hole already
> and here I attach participants' consent forms
>
> thank you so much
>
> and wish you have a good holidays
>
> Leo

4.5 Questions of the preparatory paper for the mid-term examination.

A. Filling in the Blanks (20%)

1. X follows normal distribution, its mean μ = 100 and standard deviation σ = 6. Its standard normal distribution Z =________.
2. The unemployment rate in a certain city is 16%, now 100 residents are surveyed, n = 100. Then the average of unemployment rate is ________, sample standard deviation is ________.
3. The average life of 200 cars is μ = 20, and σ = 8. If we randomly sample 36 cars, the average life is ________, and sample standard deviation is ________.

B. Calculation (80%)

1. The average life of particular fluorescent tubes is 1500 hours, and its standard deviation is 60 hours.

   (1). If someone buys 36 fluorescent tubes, what is the probability of that the average life of the 36 light bulbs is at least 1490 hours?

   (2). How many fluorescent tubes should we buy to make the probability of that the average life is more than 1490 hours equals 0.95? (10%)
2. The average life of particular lights is μ = 500 hours, and σ = 50 hours. Now we select 81 random samples, please find (1). The statistics values of the distribution of $\overline{X}$; (2). P ($\overline{X}$ < 490); (3). P ($\overline{X}$ > 500); and (4). P (485 < $\overline{X}$ < 515). (20%)
3. A particular company has 20,000 employees, and 2,000 of them are females. Now we select 100 random samples, then (1). The statistics values of the distribution of $\overline{P}$; (2). P ($\overline{P}$ < 0.04); (3). P ($\overline{P}$ > 0.16); (4). P (0.04 < $\overline{P}$ < 0.16) . (20%)

   [$\overline{P}$ represents the average of sample proportion.]
4. 600 customers are surveyed about their preference of purchasing domestic cars or purchasing imported cars? Their responses are listed below. (15%)

Preference	Domestic	Imported
Numbers	240	360

   What is the 95% confidence interval of the proportion of the customers who prefer purchasing domestic cars?
5. We sample 64 students who major in quality management in the CIT College, and find that the mean of total score of their graduation test is 510 and standard deviation is 36. Find the 90% confidence interval of the population mean. (15%)

4.6 The teacher's solutions of the preparatory paper for the mid-term examination.

A.

1. $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-100}{6}$

2. $\overline{P} = 0.16$, $\sigma_{\overline{P}} = \sqrt{\dfrac{\overline{P}(1-\overline{P})}{n}} = \sqrt{\dfrac{0.16\cdot 0.84}{100}} = 0.04\cdot\sqrt{0.84}$

3. $\overline{X} = 20$, $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{8}{\sqrt{36}} = \dfrac{8}{6}$

B.

1. (1). $\mu = 1500$, $\sigma = 60$, $n = 36$

$$P(\overline{X} \ge 1490)$$

$$= P(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \ge \frac{1490-1500}{60/\sqrt{36}})$$
$$= P(Z \ge -1)$$
$$= 1 - P(Z > 1)$$
$$= 1 - 0.1587$$
$$= 0.8413$$

[Table B users need to transfer the question format more often than Table A users do. The teacher normally used Table B in the study, and I will ask why he used Table B. He also asked the students to memorise the transferring principles, and I will examine how the approach worked.]

(2). $P(\overline{X} > 1490) = 0.95$

$$P(Z > \frac{1490-1500}{60/\sqrt{n}}) = 0.95$$

$$P(Z > \frac{-\sqrt{n}}{6}) = 0.95$$

$$1 - P(Z > \frac{\sqrt{n}}{6}) = 0.95$$

$$P(Z > \frac{\sqrt{n}}{6}) = 0.05$$

$$\frac{\sqrt{n}}{6} = 1.645$$

$$n = 36\cdot 1.645^2 = 97.4169 => 98$$

2. (1). $\mu = 500$, $\sigma = 50$, $n = 81$

$$\mu_{\overline{X}} = 500\,,\sigma_{\overline{X}} = \frac{50}{\sqrt{81}} = \frac{50}{9}$$

(2). $P(\overline{X} < 490)$

$$\begin{aligned}
&= P(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{490-500}{50/\sqrt{81}}) \\
&= P(Z < \frac{-10\times 9}{50}) \\
&= P(Z < -1.8) \\
&= P(Z > 1.8) \\
&= 0.0359
\end{aligned}$$

(3). $P(\overline{X} > 500)$

$$\begin{aligned}
&= P(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{500-500}{50/9}) \\
&= P(Z > 0) \\
&= 0.5
\end{aligned}$$

(4). $P(485 < \overline{X} < 515)$

$$\begin{aligned}
&= P(\frac{485-500}{50/9} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{515-500}{50/9}) \\
&= P(\frac{-15\times 9}{50} < Z < \frac{15\times 9}{50}) \\
&= P(-2.7 < Z < 2.7) \\
&= P(Z > -2.7) - P(Z > 2.7) \\
&= 1 - P(Z > 2.7) - P(Z > 2.7) \\
&= 1 - 2\times P(Z > 2.7) \\
&= 1 - 2\times 0.0035 \\
&= 1 - 0.007 \\
&= 0.993
\end{aligned}$$

[The teacher drew a graph when he transferred $P(-2.7 < Z < 2.7)$ to $P(Z > -2.7) - P(Z > 2.7)$ and $1 - P(Z > 2.7) - P(Z > 2.7)$ but he did not explain what the graph meant. I will examine how the participants transfer and whether they use graphs.]

3. (1). $\overline{P} = 0.1$, $\sigma_{\overline{P}} = \sqrt{\frac{0.1\cdot 0.9}{100}} = \frac{0.3}{10} = 0.03$

[The teacher did not explain what $\overline{P}$ means in the question and in his solution.]

(2). $P(\overline{P} < 0.04)$

$$= P(Z < \frac{0.04 - 0.1}{0.03})$$
$$= P(Z < \frac{-0.06}{0.03})$$
$$= P(Z < -2)$$
$$= P(Z > 2)$$
$$= 0.0228$$

[When the teacher asked the students to find the probability of Z > 2 in the table, some students interpreted .0228 in Table B as 0.228. I will notice whether the participants have such problems.]

(3). $P(\overline{P} > 0.16)$

$$= P(Z > \frac{0.16 - 0.1}{0.03})$$
$$= P(Z > \frac{0.06}{0.03})$$
$$= P(Z > 2)$$
$$= 0.0228$$

(4). $P(0.04 < \overline{P} < 0.16)$

$$= P(\frac{0.04 - 0.1}{0.03} < Z < \frac{0.16 - 0.1}{0.03})$$
$$= P(\frac{-0.06}{0.03} < Z < \frac{0.06}{0.03})$$
$$= P(-2 < Z < 2)$$
$$= P(Z > -2) - P(Z > 2)$$
$$= 1 - P(Z > 2) - P(Z > 2)$$
$$= 1 - 2P(Z > 2)$$
$$= 1 - 2 \times 0.0228$$
$$= 1 - 0.0456$$
$$= 0.9544$$

[The teacher seemed to recall the transferring principles to transfer the format. I will examine whether the students applied this approach successfully.]

4. $\overline{P} = \frac{240}{600} = 0.4$, $\sigma_{\overline{P}} = \sqrt{\frac{0.4 \times 0.6}{600}} = \frac{0.2}{10} = 0.02$

Confidence interval = point estimate ± max sampling error

$$= 0.4 \pm z_{\frac{0.05}{2}} \cdot \sigma_{\overline{P}}$$
$$= 0.4 \pm 1.96 \cdot 0.02$$
$$= 0.4 \pm 0.0392$$

[The teacher drew another graph to explain why he divided 0.05 by 2: 0.05 is the sum of two symmetric tails in the left-hand side and right-hand side, so the area of each tail equals 0.05/2 = 0.025. Then he added a different color (red) to represent the area of 90% confidence interval, and argued that the area of 90% is smaller than that of 95%. He also asked students to remember 3 frequently used value: $z_{\frac{0.1}{2}} = 1.645$ for 90% confidence interval, $z_{\frac{0.05}{2}} = 1.96$ for 95% confidence interval and $z_{\frac{0.01}{2}} = 2.575$ for 99% confidence interval. I will examine whether the participants use graphs to find z scores or directly recall the z scores.]

5. $\overline{X} = 510$, $\sigma = 36$, $\sigma_{\overline{X}} = \frac{36}{\sqrt{64}} = \frac{36}{8} = 4.5$

Confidence interval = $510 \pm z_{\frac{0.1}{2}} \cdot \sigma_{\overline{X}}$

$$= 510 \pm 1.645 \cdot 4.5$$
$$= 510 \pm 7.4025$$

[The teacher solved the question without drawing graphs or explaining how he obtained 0.1, why he divided 0.1 by 2, what $z_{\frac{0.1}{2}}$ meant and how it became 1.645. Although he had explained the process in Q4, I will examine whether the participants really understand these issues.]

4.7. Questions of the mid-term examination.

A. Filling in the Blanks (20%)

1. X follows normal distribution, its mean μ = 200 and standard deviation σ = 20. Its standard normal distribution Z =________.
2. The unemployment rate in a certain city is 5%, now 200 residents are surveyed, n = 200. Then the average of unemployment rate is ________, sample standard deviation is ________.
3. The average life of 100 cars is μ = 18, and σ = 6. If we randomly sample 36 cars, the average life is ________, and sample standard deviation is ________.

B. Calculation (80%)

1. The average life of particular fluorescent tubes is 1500 hours, and its standard deviation is 60 hours.

(1). If someone buys 36 fluorescent tubes, what is the probability of that the average life of the 36 light bulbs is at least 1490 hours?

(2). How many fluorescent tubes should we buy to make the probability of that the average life is more than 1490 hours equals 0.95? (10%)

2. The average life of particular lights is μ = 500 hours, and σ = 50 hours. Now we select 81 random samples, please find (1). The statistics values of the distribution of $\overline{X}$; (2). P ($\overline{X}$ < 490); (3). P ($\overline{X}$ > 500); and (4). P (485 < $\overline{X}$ < 515). (20%)
3. A particular company has 20,000 employees, and 2,000 of them are females. Now we select 100 random samples, then (1). The statistics values of the distribution of $\overline{P}$; (2). P ($\overline{P}$ < 0.04); (3). P ($\overline{P}$ > 0.16); (4). P (0.04 < $\overline{P}$ < 0.16) . (20%)

[$\overline{P}$ represents the average of sample proportion.]

4. 600 customers are surveyed about their preference of purchasing domestic cars or purchasing imported cars? Their responses are listed below. (15%)

Preference	Domestic	Imported
Numbers	240	360

What is the 95% confidence interval of the proportion of the customers who prefer purchasing domestic cars?

5. We sample 64 students who major in quality management in the CIT College, and find that the mean of total score of their graduation test is 510 and standard deviation is 36. Find the 90% confidence interval of the population mean. (15%)

[Most of the questions in the mid-term examination were the same to the questions in the preparatory paper except for the questions in part A.]

4.8. Questions of the preparatory paper for the final examination.

A. Filling in the Blanks (25%)

1. The population follows normal distribution and the standard deviation is known, then the test statistic of mean is ________.
2. The population follows normal distribution, the standard deviation is unknown and the sample is small, then the test statistic of mean is ________.
3. The test statistic of difference of two means of two independent normal populations with known standard deviation is ________.
4. The test statistic of difference of two means of two independent normal populations with unknown but equaled standard deviation is ________.
5. The test statistic of the difference of two population proportions is ________.

B. Calculation (75%)

1. Two groups of data are independent and follow normal distribution. The population standard deviations are unknown.

   Group A: $n_1 = 100$, $\overline{X_1} = 145$, $s_1^2 = 196$.

   Group B: $n_2 = 121$, $\overline{X_2} = 65$, $s_2^2 = 204$.

   The confidence coefficient is 99%. Find the confidence interval of the difference of the means of the two groups of data. (10%)
2. Two groups of data are independent and follow normal distribution. The population standard deviations are unknown but equal.

   Group A: $n_1 = 16$, $\overline{X_1} = 70$, $s_1^2 = 16$.

   Group B: $n_2 = 9$, $\overline{X_2} = 65$, $s_2^2 = 9$.

   The confidence coefficient is 95%. Find the confidence interval of the difference of the means of the two groups of data. (10%)
3. Department A and Department B are voting for a policy. Now we find 48 of 100 random samples from Department A support the polity whereas 30 of 80 random samples from Department B support it. Find the 99% confidence interval of the difference of proportion of supporting the policy in the two departments. (10%)
4. The lives of batteries produced by Jin-ding Company follow normal distribution.

   (1) The company announces that the average life of the battery is 12 hours and the standard deviation is 2.5 hours. The average life of 100 samples is 11.5 hours. Significant level is 5%. Test whether the company's announcement is correct.

   (2) The company announces that the average life of the battery is more than 12 hours and the standard deviation is 2.5 hours. The average life of 100 samples is

11.5 hours. Significant level is 1%. Test whether the company's announcement is correct (20%)

5. Taisho Company begins training courses. They use blackboard program in Department A and case study in Department B. They collect scores of 64 samples from Department A and 64 samples from Department B. The average score of Department A is 85 and the standard deviation is 10; the average score of Department B is 80 and the standard deviation is 8. Are the scores of the two departments significantly different? (15%)
6. One farm uses Fertiliser A and Fertiliser B on the same kind of crop. The yields are:
   Fertiliser A: 16, 13, 12, 15, 10
   Fertiliser B: 8, 7, 9, 6, 5, 7.
   Test whether the effect of the two fertilisers are the same. (Assuming the two populations follow normal distribution and the two population variances are the same. $\alpha = 0.05$)(15%)

4.9. The teacher's solutions of the preparatory paper for the final examination.

A.

1. $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$

2. $t = \dfrac{\overline{X} - \mu}{s / \sqrt{n}}$

3. $Z = \dfrac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^{\,2}}{n_1} + \dfrac{\sigma_2^{\,2}}{n_2}}}$

4. $t = \dfrac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$, and $s_p^{\,2} = \dfrac{(n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2}}{n_1 + n_2 - 2}$

5. $Z = \dfrac{(\hat{P_1} - \hat{P_2}) - (P_1 - P_2)}{\sqrt{\dfrac{P_1(1 - P_1)}{n_1} + \dfrac{P_2(1 - P_2)}{n_2}}}$

B.

Confidence interval (interval estimate)

= point estimate $\pm$ maximum error

1. $\overline{X_1} - \overline{X_2} \pm z_{0.005} \cdot \sqrt{\dfrac{s_1^{\,2}}{n_1} + \dfrac{s_2^{\,2}}{n_2}}$

$= 145 - 65 \pm 2.575 \cdot \sqrt{\dfrac{196}{100} + \dfrac{204}{121}}$

$= 80 \pm 4.916$

[The teacher did not explain how he obtained the value 0.005. He did not use the table of normal distribution to find the z score and asked the students to memorise $z_{0.005} = 2.575$.]

2. $(\overline{X_1} - \overline{X_2}) \pm t_{0.025}(16 + 9 - 2) \cdot s_p \cdot \sqrt{\dfrac{1}{16} + \dfrac{1}{9}}$

$= (70 - 65) \pm 2.069 \cdot 3.6831 \cdot \sqrt{\dfrac{1}{16} + \dfrac{1}{9}}$

$= 5 \pm 3.175$

[The teacher asked the students to use calculators to obtain

$$s_p{}^2 = \frac{(n_1-1)s_1{}^2+(n_2-1)s_2{}^2}{n_1+n_2-2} = \frac{(16-1)16+(9-1)9}{16+9-2} = \frac{15\cdot16+8\cdot9}{23} = \frac{240+72}{23} = \frac{312}{23}$$ and

$s_p = \sqrt{\frac{312}{23}} = 3.683$ for him.]

3. $(P_1-P_2)\pm z_{0.005}\cdot\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}$

$$= (0.48-0.375)\pm 2.575\cdot\sqrt{\frac{0.48\cdot0.52}{100}+\frac{0.375(0.625)}{80}}$$
$$= 0.105\pm0.19$$
$$= (-0.085, 0.295)$$

4.

(1) $\begin{array}{l} H_0: \mu = 12 \\ H_1: \mu \neq 12 \end{array}$

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

$$= \frac{11.5-12}{2.5/\sqrt{100}}$$
$$= \frac{-5}{2.5} = -2 < -z_{0.025} = -1.96$$

$\therefore$ reject H_0, the average life is not 12 hours.

(2) $\begin{array}{l} H_0: \mu \geq 12 \\ H_1: \mu < 12 \end{array}$

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = -2 \not< -z_{0.01} = -2.325$$

Accept H_0, the average life is more than 12 hours.

5. $\begin{array}{l} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{array}$

$$Z = \frac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{85 - 80}{\sqrt{\frac{10^2}{64} + \frac{8^2}{64}}} = \frac{5 \times 8}{\sqrt{164}} > z_{0.025} = 1.96$$

$\therefore$ reject H_0, the scores are significantly different.

6. $n_1 = 5, \overline{X_1} = 13.2, s_1^2 = \frac{\sum (x_i - 13.2)^2}{4} = 5.7$

$n_2 = 6, \overline{X_2} = 7, s_2^2 = \frac{\sum (x_i - 7)^2}{5} = 2$

$\therefore s_p^2 = \frac{4s_1^2 + 5s_2^2}{9} = \frac{4 \cdot 5.7 + 5 \cdot 2}{9} = \frac{32.8}{9}, s_p = \frac{5.5}{3} \doteqdot 1.9$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$$t = \frac{13.2 - 7}{1.9 \cdot \sqrt{\frac{1}{5} + \frac{1}{6}}} = \frac{6.2 \cdot \sqrt{30}}{1.9 \cdot \sqrt{11}} = 5.389 > t_{0.025}(9) = 2.262$$

$\therefore$ reject H_0, the effects are significantly different.

[All of the preparatory questions required using formulas and calculations. The teacher did not draw any graph in his solutions. He also did not explain how he obtained the z scores according to the confidence coefficient, significant level and α. I will address on these issue in the clinical interviews.]

4.10. Questions of the final examination.

$$Z_{0.975}=1.96, Z_{0.99}=2.325, Z_{0.995}=2.575,$$
$$t_{0.01}(10)=2.764, t_{0.01}(11)=2.718, t_{0.1}(11)=1.363$$
$$S_p^{\ 2}=\frac{(n_1-1)S_1^{\ 2}+(n_2-1)S_2^{\ 2}}{n_1+n_2-2}$$

A. Filling in the Blanks (25%)

1. The population follows normal distribution and the standard deviation is known, then the test statistic of mean is ________.
2. The population follows normal distribution, the standard deviation is unknown and the sample is small, then the test statistic of mean is ________.
3. The test statistic of difference of two means of two independent normal populations with known standard deviation is ________.
4. The test statistic of difference of two means of two independent normal populations with unknown but equaled standard deviation is ________.
5. The test statistic of the difference of two population proportions is ________.

B. Calculation (75%)

1. Two groups of data are independent and follow normal distribution. The population standard deviations are unknown.

   Group A: $n_1=100$, $\overline{X_1}=145$, $s_1^{\ 2}=196$.

   Group B: $n_2=121$, $\overline{X_2}=65$, $s_2^{\ 2}=204$.

   The confidence coefficient is 99%. Find the confidence interval of the difference of the means of the two groups of data. (10%)
2. Two groups of data are independent and follow normal distribution. The population standard deviations are unknown but equal.

   Group A: $n_1=16$, $\overline{X_1}=70$, $s_1^{\ 2}=16$.

   Group B: $n_2=9$, $\overline{X_2}=65$, $s_2^{\ 2}=9$.

   The confidence coefficient is 95%. Find the confidence interval of the difference of the means of the two groups of data. (10%)
3. Department A and Department B are voting for a policy. Now we find 48 of 100 random samples from Department A support the polity whereas 30 of 80 random samples from Department B support it. Find the 99% confidence interval of the

difference of proportion of supporting the policy in the two departments. (10%)

4. The lives of batteries produced by Jin-ding Company follow normal distribution.

(1) The company announces that the average life of the battery is 12 hours and the standard deviation is 2.5 hours. The average life of 100 samples is 11.5 hours. Significant level is 5%. Test whether the company's announcement is correct.

(2) The company announces that the average life of the battery is more than 12 hours and the standard deviation is 2.5 hours. The average life of 100 samples is 11.5 hours. Significant level is 1%. Test whether the company's announcement is correct (20%)

5. Taisho Company begins training courses. They use blackboard program in Department A and case study in Department B. They collect scores of 64 samples from Department A and 64 samples from Department B. The average score of Department A is 85 and the standard deviation is 10; the average score of Department B is 80 and the standard deviation is 8. Are the scores of the two departments significantly different? (15%)

6. The length of films produced by two companies are:

   Company A: 102, 98, 109, 86, 92

   Company B: 81, 97, 134, 92, 165, 87, 114

   Assuming the lengths of the films follow normal distribution and the two population variances are the same. $\alpha = 0.05$. Test whether the length of film produced by company B 10 minutes longer than the length of film produced by Company A. (15%)

$$Z_{0.975} = 1.96, Z_{0.99} = 2.325, Z_{0.995} = 2.575,$$
$$t_{0.01}(10) = 2.764, t_{0.01}(11) = 2.718, t_{0.1}(11) = 1.363$$
$$S_p^{\ 2} = \frac{(n_1 - 1)S_1^{\ 2} + (n_2 - 1)S_2^{\ 2}}{n_1 + n_2 - 2}$$

[The teacher attached three z scores, three t scores and the formula of s_{p^2} at the beginning and the end of the question sheet. With the attachment, the students did not need to use the table of normal distribution and the table of t distribution. In the student interviews, I notice that the students had difficulties to use the tables of distribution. Most of the questions in the final examination were the same to the questions in the preparatory paper except for the sixth question of calculation.]

6.1. Questions of the first teacher interview.

Part I

Q1. What is the meaning/advantage of visualisation in your statistics learning?

Q2. What is the meaning/advantage of visualisation in your statistics teaching?

Q3. From the beginning to the chapter of estimation, in which section do you feel visualisation the most needed?

Q4. What aspects of visualisation help dealing with, explaining or solving better than non-visual (symbolic) representation?

Part II

Q5. Why do you choose this table of standard normal distribution (Table B)?

Q6. Does Table B have any advantage that other tables do not have?

Q7. In your explanation of one question in the preparation question last week:

$$\begin{aligned} & P(-2.7 < Z < 2.7) \\ & = P(Z > -2.7) - P(Z > 2.7) \\ & = 1 - P(P > 2.7) - P(Z > 2.7) \end{aligned}$$

Why do you choose to recall the principle rather than to use graph?

Part III

Q8. You asked the students to memorise the three z_α scores corresponding to 90%, 95% and 99% confidence intervals. Why do you ask them to memorise rather than find them in the Z tables? Did you find they had difficulties in learning/realising this?

Q9. As mentioned in Q8, you drew a graph to represent what 0.95% and 0.90 mean and that 0.95 is broader than 0.90. How do your students realise this? Also, do they understand what confidence interval is?

Q10. When you asked the students what value should be used in the formula of standardisation, many of them were silent. Do they understand what standardisation is?

Q11. When teaching standardisation, did you use a graph to help students learn this? If yes, why and how? If not, why not and which method did you use?

6.2. The first teacher interview on 23rd Apr 2009.

L (Leo, the interviewer): Hello Mr. A, thank you for being my participant. There are a few things I have to know before the interview. First, how many students are in this statistics course?

TA (Teacher A): Fifty-one, including forty-nine Year two students and two Year four students.

L: Have you taught any statistics course before this academic year?

TA: Yes, but only once in the department of management information (this department was industry management a few years ago) from September 07 to June 08.

L: Did you also use this book last time?

TA: No, this book was chosen by the previous teacher in the first semester in this academic year. I replaced the teacher in the second semester but I did not change textbook.

L: Alright, thanks. Let’s begin with the interview questions.

PART I

Q1: What is the meaning/advantage of visualisation in your statistics learning?

TA: When I studied in my bachelors’ degree of mathematics, the probability course was arranged in the first semester of Year two, and statistics was arranged in the second semester of the same year. However, I did not learn statistics well because my teachers in the two subjects did not link them well and there was a gap between them.

L: How about visualisation in your learning?

TA: There was not visualisation because my teacher taught us numerical statistics. I learned statistics again in my PhD degree two years ago, but the teacher did not use graphs often, except for drawing normal distribution, t distribution, and using the tables.

Leo: When I taught statistics a few years ago, I might be too mathematised as I paid most attention to the formulas and calculation in problem-solving. When I observed your explanation of mid-term preparation, an idea just came to my mind: many of your questions involved normalisation, and do you think your students really know the meaning of normalisation?

TA: No. I did not emphasise it, and students are only required to know how to normalise and get the answer.

Q2. What is the meaning/advantage of visualisation in your statistics teaching?

TA: Even I learned statistics well in PhD degree, I still used little graph.

L: How did you learn well?

TA: I memorised them, but now they are forgotten. My teacher majored in statistics, so he emphasised deduction or transformation, such as how t distribution is transformed to normal distribution.

L: In class, do you use visualisation?

TA: In fact, I prefer to use simple examples, such as to lift a watermelon and guess how heavy it is: the idea of estimation. Besides, it is not possible to be so exact, and there will be some error: the idea of interval. Then I taught students to use different formulas for large samples and small samples. For the usage of visualisation, I drew a graph when looking for probability in the tables. For instance, I can use the graph to explain and help students understand why P $(t > 1)$ and P $(t < -1)$ are the same by using the property symmetry. I told my students that you can draw the graph if you have any problem in using the table. However, I still do not like to use the software PowerPoint.

L: In my research, visualisation includes not only PowerPoint, but also other forms such as drawn pictures on paper. When I look for probability in tables, I always draw graphs because the area of graphs can correct/check the probability value, such as a large area in graph and a small probability in the table which may be caused by ignoring the negative symbol.

Q3. In what situations from the beginning of the course to the chapter of estimation, is visualisation needed the most by you?

L: There is one thing to notice: when we discuss statistics here, probability is not included.

TA: Students cannot understand normal distribution without the graph. Taking the table of normal distribution I use in this class for example, students cannot understand what the values in the table mean if they were not taught the right-hand tail graph.

L: Did you teach them t distribution, F distribution and chi-square distribution?

TA: Yes, I taught them before sampling.

Q4. In what aspects does visualisation help deal with, explain, or solve better than non-visual methods?

TA: This question is similar to Q3. However, I think there are not many places where graphs are necessary.

L: Do you allow students to use their books and notes in the mid-term?

TA: No, but I gave them mid-term preparation questions.

L: Why are eighty percent of questions totally the same as the preparation questions?

TA: Students' levels are low (laugh).

L: How about last semester which was taught by another teacher?

TA: The teacher did not give them preparation questions, but he gave them homework to practise and answers to check.

L: In your questions, there is nothing about t, F or chi-square distribution. May I ask why?

TA: In t or F distribution, calculation is necessary but it is very troublesome for students. Actually I like to use t distribution in questions, but even if n (sample numbers) is small, their calculating abilities are still too bad.

L: When I taught statistics a few years ago, I also gave my students preparation questions. Most of the questions in preparation are similar to those in the mid-term and final; numbers were changed but conditions and principles were the same. Sometimes I changed the order of the questions, and students just followed the process in the preparation questions. I think they do not understand what distribution or formula should be used in each condition.

PART II

Q5. Why do you choose this table of standard normal distribution (Table B)?

L: There are 2 types of normal table and graph (A and B) in this textbook, why do you use Table B?

TA: Table B is more often used. I read many books, and most of them use Table B which is right-hand side tail.

L: I use Table A. We cannot say which one is definitely better or worse than the other, but Table B does not include the negative part. Do you think this may cause your students difficulty?

TA: Yes, so I have to teach them how to transform by using symmetry.

L: Do they have difficulty in the process of transforming?

TA laughed: I don't know. I haven't marked their mid-term paper yet.

L: There is another kind of normal table and graph often used: shadowed area between 0 and z score. I wonder whether the type of normal table and graph chosen by teacher influence students' learning of normal distribution, and how?

TA laughed but did not answer.

Q6. Does Table B have any advantage that Table A and C do not have?

L: Can you tell me what makes you choose Table B?

TA: Because my teachers used this, and I am used to it. I never compared or considered differences between these types.

L: Do you know which type was used by the teacher in the first semester?

TA: No, I did not ask.

L: That's fine. I can ask them in my interview next week.

TA: OK, then you can tell me. If the students were taught another type, you can ask them which type fits them better.

L: In this book, Table C is not used. I thought this type is strange.

TA: I agree with you. It is really strange.

L: However, Table C has its own advantage: it is convenient when looking for confidence intervals. For example, when we look for 95% confidence interval, we only have to divide 0.95 by 2, and get 0.475, then find 0.475 in the table without using the idea of α, 1 - α, α / 2 and 1 - α / 2.

TA: That is right! I never considered this before.

Q7. In your explanation last week: P (-2.7 < Z < 2.7) = P (Z > -2.7) – P (Z > 2.7) = 1- P (Z > 2.7) – P (Z > 2.7), why did you choose to remember principles, rather than to use graphs?

L: In fact, each of the three types has its own advantage and its own weakness. I was interested by your answering process that

$$P(-2.7 < Z < 2.7) = P(Z > -2.7) - P(Z > 2.7) = 1 - P(Z > 2.7) - P(Z > 2.7).$$

I saw you drew a picture to explain this, but you stressed the principles and asked students to remember the principles, rather than to understand by using graphs, why?

TA: In the previous class, I drew two pictures and used the idea of area that the middle area equals the large area subtracted by the small area. I do not know how many students drew graphs in mid-term papers yet, but I will notice this when I mark them.

L: Thank you very much. Perhaps I can find some students who rely on graphs by doing this and ask them to participate in my research.

TA: I hope so. But I don't think they will draw a graph since they have the preparation paper.

L: That is fine. Did you have common test in class?

TA: Yes, I test them in order to have the opportunity of evaluating their performance.

Q8. You also asked students to remember the three z_α values and their corresponding confidence intervals, why do you ask them to remember? Did you find they have difficulty in learning/realising this?

L: In your explanation of mid-term preparation, you asked students to memorise the corresponding z score and (1-α) confidence interval. Did you tell them how to find the z value by using the table?

TA: Yes, I taught them how to do it. I asked them to remember these numbers because it is more convenient for students, and beneficial for their grades. Also, I don't have to worry about their poor score that much. I asked the students to remember the numbers corresponding to confidence intervals because I planned not to give them the table of normal distribution in the mid-term examination. It is difficult and troublesome. However, I still gave the table.

L: You don't want to test their ability of looking for probabilities in the table, because you think they have problems here?

TA: Yes, I already tested this in class, but they could not do it.

Q9. As Q8, you drew a graph to represent what 0.95 and 0.90 mean and that 0.95 is broader than 0.90, how do your students realise this? Also, do they understand what confidence interval is?

L: Do you think your students understand that 95% confidence interval is wider than 90% confidence interval?

TA: Yes, they do. I taught them that this is the idea of area, and 0.95 is bigger than 0.9.

L: Do you think they understand what confidence interval really is?

TA: It is not easy to say, and I think they do not really know. I told them to consider how big a difference (error) they can accept, and that the interval is wider when they can accept bigger difference. I don't know how much they understand, and especially statistics is very difficult for new learners.

Q10. When you asked students what value should be used in the formula of

normalising (standardising), many of them are silent. Do they know what normalising is? When teaching normalising, did you use graphs to help students learning this? If yes, why and how? If not, why not and what is your method?

L: In your preparation paper, normalisation is used in many questions, but students did not answer you when you asked them what numbers should be used in the normalising process. Therefore, I need to ask: do you think your students understand what normalisation means?

TA: I did not emphasise this. I just told them the goal of normalisation is to look for probability in the table. I think they do not understand the real meaning of normalising. In my class, I prefer to teach my students basic ideas by daily-life examples. I am glad to participate in your research which may help me know what students are thinking.

L: Thank you very much, and see you next week.

TA: Bye.

6.3. Questions of the first student interview.

Part I

Q1. Please look for the value in the table of normal distribution. P (Z < -0.03), P (Z < 1.22), P (Z > -3.1), P (Z > 2.05), P (-1.4 < Z < 2.32), P (-2.41 < Z < -0.03), and P (Z = 3).

Q2. How did you realise the principles? What helped you in the process?

Q3: Please look for $t_{0.005}(9)$ and $t_{0.99}(23)$.

Q4. How did you realise the principles? What helped you in the process?

Part II

Q5. How do you define visualisation? What is included in your definition of visualisation?

Q6. Did you use visualisation in your statistics learning? If yes, how and why? If no, why not?

Q7. Did you use visualisation in the mid-term? If yes, does it help and how? If not, why not?

Q8. From the beginning to the chapter of estimation, which section do you feel visualisation is needed the most for you?

Q9. Do you find/feel any function/advantage of visualisation that non-visual methods cannot serve or cannot serve so well?

Part III

Q10. What type of table and graph of standard normal distribution did you learn in the first semester? Did the teacher in the first semester use type B as well? Do you know other type(s)?

Q11. If you know more than one type, which one do you prefer and why? How well do you understand the types? Can you compare their strength and weakness?

Q12. Did you have difficulty in using the table? If yes, describe it.

Part IV

Q13. Could you please explain your understanding of the process of standardisation (normalisation)?

Q14. Could you please explain your understanding of the process of standardisation (normalisation) by graph?

Q15. Do you remember the three z_{α} scores corresponding to 90%, 95% and 99% confidence intervals? Do you know how to get them? Can you explain the process?

Q16. Do you know what confidence interval means? What is the difference between 0.95 and 0.90 confidence intervals? Explain in your own language.

6.4. The first student interview with S1 and S2 on 29th Apr 2009.

PART I

Before the interview, S1 and S2 were given questions to answer:

Q1. Please look for probability in the table of normal distribution. P (Z < -0.03), P (Z < 1.22), P (Z > -3.1), P (Z > 2.05), P (-1.4 < Z < 2.32), P (-2.41 < Z < -0.03), and P (Z = 3).

After 12 minutes, they said that they had finished.

L: Would you like to compare your answers? Some of them are different, and it means some of them are wrong.

S2: Then the wrong one should be me.

S2 asks S1: Should I write “1 –” here?

S1: Yes.

S2: But I didn’t.

S1: You have to see whether it is > or <, and it’s the main point.

S2: I know.

S2: I think I am wrong, because I did not concentrate in class.

L: Let’s see how many answers are different?

	P (Z < -0.03)	P (Z < 1.22)	P (Z > -3.1)	P (Z > 2.05)	P (-1.4 < Z < 2.32)	P (-2.41 < Z < -0.03)	P (Z = 3)
Correct Answer	0.488	0.8888	0.999	0.0202	0.9898 - 0.0808 = 0.909	0.488 – 0.008 = 0.48	0
S1 (prefers A but only got B)	P (Z < -0.03) = 0.5120	P (Z < 1.22) = 0.8888	P (Z > 3.1) = 0.0010	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) ↓ =0.5667 0.898 =0.3231	P (-2.41 < Z < -0.03) = 0.504	P (Z = 3) = 0.5
S2 (prefers type B)	-0.03; 0.4880	< 1.22; 0.1112	> - 3.1; 0.010	> 2.05; 0.202	-1.4 < Z < 2.32; 0.0808 < Z < 0.0102	-2.41 < Z < -0.03; 0.0080 < Z < 0.4880	

S1: Only 2 are the same.

S2: Ha hah.

L: Really?

S2: Because I didn’t subtract,… I don’t know how to deal with that.

L: Is it?

S2: Yes.

L: Did your teacher teach that?

S1 and S2: Yes.

L: How about the last question?

S2: I didn’t answer it.

L: Do you feel the question was difficult?

S1: I feel this is ok.

S2: I can only answer the positive questions.

L: Why only positive?

S2: Because I cannot subtract.

L: Did your teacher teach that?

S2: Yes, but I did not learn well.

L: Ok, let's see the answers.

L: Some questions are designed to see how you may make a mistake, such as 0.202.

S2: Oh, I just find that how to deal with the negative part, I should find the probability in the table, then 1 – that probability. Ok, I understand now.

L: These questions include Z smaller than a value, Z bigger then a value, and Z between 2 values.

S2: I cannot do questions of the Z between 2 values.

L: Can you do this kind of question, S1?

S1: Yes, I can.

L: How?

S1: Use the table to find two probabilities and subtract to get the answer of middle area.

[Some of students' own language is difficult to understand].

L: Why subtract?

S1: The two found values are about the position, and I subtract to get the middle area.

L: You can find the value by using the table.

S1: Yes.

L: Then why do you draw a graph when answering P (Z = 3)?

S1: The *equal* question is not often mentioned, and in my memory, I think I should take the middle value, like when z = 0. There is no > or <, so it is difficult to *guess* which side the *frequency* is.

L: Oh, did teacher A teach you that?

S1: I think he normally mentioned > or <. When I begin solving this question, I think *bigger than* subtracts *smaller than* becomes *equal*.

S2: It's not.

S1: I need some time to think about it.

L: Does the graph you just drew help?

S1 laughed: No, it's useless.

[S1 did not realise the probability is the area].

L: What does your graph mean?

S1: One side is bigger than, the other side is smaller than, and the middle is equal to. I think the middle equal should be zero. Each of the two sides is 50%.

L: Can you write down where 3 is in your graph?

S1: 3 is in the middle. It does not change, and it equals 0.

L: In your graph, is the middle 3?

S1: Yes.

L: Why did you write 0 below the middle?

S1: Because I looked at it in the beginning. If 3 is in the left side or right side, I think the answer may be 0.03 or 0.05, but since the middle one is the middle column, I think both of each side is a half. It's just my personal idea.

L: S2, you had an answer for each question except for the last one, did you do that by using the table?

S2: Yes.

L: For the 5^{th} and 6^{th} questions, are they your result?

S2: Yes, *this subtracts that to get the middle*.

L: But it's a negative number.

S2: Hmm… I …

L: The bigger one subtracts the smaller one?
S2: No, I don't know how to continue.

Leo: Ok, let' try t distribution. Is it taught?
S1: Yes.

Q3: Look for t $_{(0.005, 9)}$, t $_{(0.99, 23)}$.
S1: Oh, I just found a wrong one.
[One minute later]
S1: Oh, wrong again.
S2: I cannot find the second one.
L: That's fine, just write it down. But I find that you don't know what to do if you cannot directly find the answer in the table.
S2: That's correct.

	t (0.005, 9)	t (0.99, 23)
Correct answer	3.25	-2.5
S1	T (0.005) $_9$ = ~~6.594~~ ~~4.781~~ 3.250	t (0.99) $_{23}$ = – 2.5000
S2	t (0.005, 9) = 3.250	t (0.99, 23) cannot find

Q2 & Q4. How did you realise the principles? What helped you in the process?
L: I noticed that teacher A asked you to remember some principles to deal with the questions whose answer cannot be found directly in the table.
S1: He asked us to remember the numbers corresponding to *interval confidence*; he also told us to use symmetry to think of *how to write the other side*. He did not teach difficult material.
S2: I did not concentrate in class.
S1: I think I won't have to use this table to do business.
L: Yes, you can use computer software to get the values, but you should understand what it means.

PART II
L: As I mentioned about my research in the beginning, I am interested in your ideas.
Q5. How do you define visualisation? What is included in your definition of visualisation?
S1: I feel that the drawn pictures make the questions more easily understood, because they show which side it is. Sometimes I am confused about bigger than and smaller than. In fact, the point is obvious and you just need to know how to deal with bigger than and smaller than.
L: Besides the graphs beyond the table of probability, what else are included in visualisation? You can use your textbook.
S1: Bar graph, point graph.
L: Do you have other answers?
S2: No.

Q6. Did you use visualisation in your statistics learning? If yes, how and why? If no, why not?

S1: I did not draw graphs often.

L: But you just did.

S2: I don't know how to draw it.

L: How do you see graphs?

S2: Graphs make it easier to understand. You can mark the number on the graph and it becomes clear.

L: Why did you not draw?

S2: I did not concentrate in class. Drawing a graph makes it easier for me to think of the answer, and less likely to be wrong.

Q7. Did you use visualisation in mid-term? If yes, does it help and how? If not, why not?

S1: No.

S2: No.

L: Why not?

S1: I didn't draw even if I did the practice questions in the textbook. I will use the graph in the front content, and see how the formula is written.

L: You didn't draw, but you still use the graph in the textbook?

S1: Yes. I will use it.

S2: But I cannot understand the graphs.

L: Are you allowed to use the book in the mid-term?

S1 and S2: No, but we were given the table.

S2: But this table is useless. Many students *memorise* the answers of the exam paper. I also memorise the formula of the question in the part of filling in the gap.

Q8. In what situations from the beginning of the course to the chapter of estimation, is visualisation needed the most by you?

S1: The teacher (teacher C) in the first semester used both PowerPoint and blackboard in class. I did not concentrate.

L: What is the content in his PowerPoint? Is it the same as the textbook?

S1: Yes, they are the same. We do practice questions by ourselves and get the answers from teacher C. Personally I don't draw any graph. I use the examples in textbook to find how to solve the questions, or ask others.

S2: I use the examples in textbook to see how they solve the questions, and follow their steps.

Q9. Do you find/feel any function/advantage of visualisation that non-visual methods cannot serve or cannot serve so well?

S1: A graph makes it clear, unlike purely numbers which are chaotic. Taking bar graph for instance, the positions can represent the height and time. When the question is with the graphs, the numbers are not so chaotic, and I can better know which number to put in which place.

L: Do you have any opinion?

S2: No.

L: I will give you the list of questions in this interview next week, and you can reply to me at anytime if you have new thoughts.

S1 and S2: Ok.

PART III

Q10. What type of table and graph of standard normal distribution did you learn in the first semester? Did teacher C use type B as well as teacher A? Do you know other type(s)?

S1: Both tables are taught by teacher A, but I am not sure which was/were taught in the first semester.

S2: I don't remember, either.

L: Did you learn statistics before this academic year?

S1 and S2: No.

L: Do you know other types?

S1 and S2: No.

Q11. If you know more than one type, which one do you prefer and why? How well do you understand the types? Can you compare their strength and weakness?

L: Which type do you prefer or which do you use more often? Type A or type B?

S1: I prefer the table *with a large area* (type A).

S2: I prefer type B.

[S1 and S2 did not bring textbooks. I photocopied Table B for S1, and let S2 use my textbook in this interview.]

L: I am sorry, I should also give you type A.

S1: Actually the two types are the same.

L: Why do you prefer type A?

S1: Because the area is big. The small area may confuse me.

L: Which type do you feel easier/more difficult to use?

S1: No difference.

L: In which situation, do you feel that you should use type A? And when should I use type B?

S1: Type A is easier to look for negative z, because the first half is about negative z. All negative z scores are over there.

L: S2?

S2: I am not good in negative z.

L: How about P (Z < 1.22)? There is no negative z.

S2: 1- the found z score and I can get the answer of negative z.

L: But in P (Z < -0.03), - 0.03 is negative, why didn't you write 1 - ?

[S2 did not notice that the principles may be influenced by > or <.]

S2: I am used to it, it is easier to use. But I feel negative z looks strange.

Q12. Did you have difficulty in using the table? If yes, describe it.

L: When do you think your preferred type of table is not easy for finding the answer?

S2: There are differences between positive and negative, and > and <, I can use the table when there is negative, and transfer < to >.

L: If you only use one kind of table, do you have difficulty using it?
S2: If the question asks a positive, and I look at the table of negative, I don't know how to find it.
L: That's because you use the wrong one.
S1 and S2 laugh.
L: I feel that there is a problem with the normal table in your textbook. The negative half and the positive half are on the other side of a piece of paper, and the user cannot see the whole table at the same time. Users may forget to turn the page.
S2: Oh, yes.

PART IV

Q13. Could you please explain your understanding of the process of normalisation?
L: Do you know what normalisation means? You look confused.
S2: No.
S1: I may have an impression.
L: It is tested in mid-term.
S2: Which question?
L: Perhaps teacher A did not emphasise the noun normalisation. It is the process of normal random variable X to standard normal variable Z. So what is your understanding of normalisation?
S1: It is only a formula for me (laugh).
L: For what?
S2: Transfer to Z.
L: Then?
S1: To get a fixed unit. It can be used in business or cost.

Q14. Could you please explain your understanding of the process of normalisation by graph?
S2: No, I cannot.
S1: Normalisation, is it a straight line, and the average in the middle? Differences are in both sides, and *all things in the range are standard*. Hmmm, it sounds strange. I should say that the area is the sum, and the middle point is the *summary average*. It's still strange.
L: You can use your book, did teacher A draw the graph?
S1: Yes, he did, but I forget how to draw it.
L: That's fine.

Q15. Do you remember the three z_α scores and their corresponding confidence intervals, do you know how to get them? Can you explain the process?
L: Can you find the z score corresponding to 95% confidence interval in the normal table by yourself?
S1: Is it in this table?
L: That's the table of t distribution.
S1: 0.95 is 005, and 0.05 divided by 2 is *0.0025*.
[S1 cannot find 0.0025 in the table.]
S1: I forget…

L: In fact, you just mentioned the right idea.
S1 (recall): *0.96* corresponds to 95%, and 99% correspond to *2.757*.
L: Can you find the z score corresponding to 99%? You should write it down.
S1: Is it the table? (HE POINTS AT TABLE A)
L: Actually you point at zero point…
S2: 2.5
L: 2.5?
S1: Yes.
L: Right side of 2.5, 0.9949 and 0.9951, but you didn't find the answer, did you?
S1: Because I was confused.
L: In type B, you remembered two-point-five-something-something, and you found 0.0051 and 0.0049, but you did not find the answer and just left.
S1: If I subtract these two number…
L: Why subtract?
S1: Let me think, no, it should be divided by two.
L: You recall it now?
S1: It may be already taught.
L: But you were asked to remember them.
S1: It is easy to remember, so I did not think of how it is found.

Q16. Do you know what confidence interval means? What is the difference between 0.95 and 0.90 confidence intervals? Explain in your language.
S2: Teacher A mentioned it very often, but I never remember.
S1 laughs
S1: It is used to compare products, such as dipper. If the difference is not *between zero*, there is a difference.
L: What is the difference between 95% confidence interval and 90% confidence interval?
S1: One is 1.645 and the other is 1.96.
L: Then how does it influence confidence interval?
S1: It will cause different numbers.
L: Which one is bigger?
S1: Perhaps, 95% may be *bought by more people*.
L: More people buy it?
S1 laughed: I also think so. It will be *more precise*.
S2: Because it is closer to 100%.
L: You think 95% is more precise than 90%?
S1: Yes. I did not listen carefully in class (laugh).
L: In fact, it is not more precise but wider ranged.
S1: Wider range? Have more choices?
L: Such as how heavy am I? When I say I am between 75 and 85, or between 70 and 90, which one do you think is more precise?
S1 and S2: Hmmm.
L: Please give me the paper you just wrote in this interview. Before you go, let's see how you answered the questions.

[S1 and S2 compared their answers to the correct ones]

Q1
S1: I was wrong in question 1.

Q3
L: Z > -3.1, you seem always to have difficulty when there is a negative.
S2: Hah.
L: I think teacher A told you that P (Z > -3.1) = 1 – P (Z > 3.1). There are some principles.
S1: Is it negative? Oh, I didn't see it.
L: Yes, you ignored it.

Q4
L: P (Z > 2.05)
S2: I was wrong, a zero is missing.

Q5
L: 2.32, should be 0.9898, and you wrote 0.898
S1: What?! What did I do?

Q7
L: It is "="
S1: Oh!!! That's it!!!

L: Thank you so much for this interview, our next interview will be 3 weeks later. See you later.
S1 and S2: Bye.

[This interview took much longer (about 70 minutes) than I thought (35-40 minutes), and I apologised for the delay. S1 and S2 did not show negative attitude for this. I told other interviewees that the interview might take more than one hour, and they accepted that. The interviews with S3 and S4, S5, and group of S6, S7, S8, S9 and S10 also took about 70-80 minutes.]

6.5. The first student interview with S3 and S4 on 29th Apr 2009.

PART I

Before the interview, S3 and S4 were given questions to answer:

Q1. Please look for probability in the table of normal distribution. P (Z < -0.03), P (Z < 1.22), P (Z > -3.1), P (Z > 2.05), P (-1.4 < Z < 2.32), P (-2.41 < Z < -0.03), and P (Z = 3).

After 13 minutes, they stopped.

L: Ok, let’s see the answers. S3, I find that most of your thoughts are correct, but you made mistakes in calculation and found probabilities which are next/close to the correct one. For example, P (Z > 2.32) = 0.0102, but you use *0.0104* which is next to 0.0102.

S3 laughed.

L: Your answer to question 5 was different to mine.

S3: I subtract this by that…

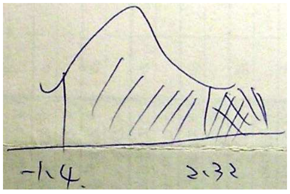
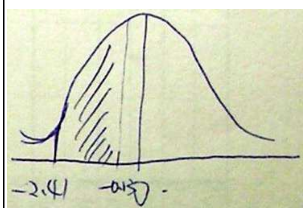
L: I think it’s an incorrect calculation. The process of calculation is not my interest, and I am more interested in how you find the values. There are other kinds of mistake. Such as in question 6, you write P (Z > -0.03) as P (Z > -0.13).

S3 and S4 laughed.

L: That’s ok. I am more concerned with your methods.

S4: I *memorise* the answers.

L: That’s fine.

	P(Z<-0.03)	P(Z<1.22)	P(Z>-3.1)	P(Z>2.05)	P(-1.4<Z<2.32)	P(-2.41<Z<-0.03)	P(Z=3)
Correct Answer	0.488	0.8888	0.999	0.0202	0.9898 - 0.0808 = 0.909	0.488 – 0.008 = 0.48	0
S3 Type B	P(Z<-0.03) = 0.4880	P(Z<1.22) = 0.8888	P(Z>-3.1) =1-P(Z>3.1) =1-0.0010 =0.999	P(Z>2.05) = 0.0202	P (-1.4 < Z < 2.32) =P(Z>-1.4)-P(Z>2.32)  =1-P(Z>1.4)-P(Z>2.32) =1-0.0808-0.0104 =*0.088*	P(-2.41<Z<-0.03)  =P(Z>-2.41)-P(Z>*-0.13*) =1-P(Z>2.41)-[1-P(Z>0.13)] =Z>0.13-P(Z>2.41) =0.4483-0.0080 =0.4403	P(Z=3) = *0.0013* *or* ≒ 0
S4 Type A	P(Z<-0.03) = 0.5120	P(Z<1.22)	P(Z>-3.1) =0.0010	P(Z>2.05) = 0.9798	P (-1.4<Z<2.32)	P(-2.41<Z< -0.03)	P(Z=3) = 0.9987

Q2. How did you realise the principles? What helped you in the process?

L: For instance, why do you transfer P (Z > -2.41) to 1 – P (Z > 2.41)?

S3: I take advantage of graph. 0 is in the middle, -2.41 is smaller than 0, and -0.03 is a little bit left side of 0. The answer is the shadow area. This area can be gotten from the difference of the whole area subtracted by this area and the whole area subtracted by that area [1 – P (Z > 2.41)] – [1 - P (Z > 0.13)].

L: Can you say again? And replace this and that by numbers.

S3: In question 5, the whole area equals 1. We cannot know how much P (Z < -1.4) is, but we can find how much P (Z > -1.4) is.

L: How?

S3: There is no negative in the table of type B, so we can transfer > to < or 1 – P (Z > +).

L: Do you always use type B?

S3: Yes.

L: Why don’t you use type A?

S3: Because teacher A emphasised type B.

L: Did he teach you both types?

S3: Yes, he taught us both types, but he emphasised type B.

S3: P (Z > 2.32) is the right side, and it is about 0.0104. Then I subtract 1 – P (Z >1.4) by 0.0104.

L: I see. S4, do you understand these principles?

S4: Yes, I do.

L: So what is your problem? You did not answer all questions.

S4: I don’t understand the formulas.

L: How about the table?

S4: I should be able to use it.

L: In question 4, can you show me how to find it?

S4: 0.0202

L: Why did you write 0.9798 ten minutes ago?

S4 laughed: Because I used type A.

L: In type A, do you just look for 2.05 and ignore smaller than or bigger than?

S4: Yes.

L: In question 1, which type do you use?

S4: Type A.

L: Where is 0.512?

S4 pointed at 0.003: Here.

L: Did you not notice that it is negative?

S4: No.

L: If you look at the page of negative z, you may find the answer.

S4: Hmm.

L: You did not answer question 2. In question3, how did you find P (Z > -3.1)? Did you directly use the table?

S4: Yes.

L: Which table?

S4: Type A.

L: Where is -3.1?

S4 pointed at: Here.

L: Is it 0.001?

S4: Hmmm.

L: You just wrote down 0.001 as the answer, and you did not notice that.

S4 laughed: Negative or positive.

L: Oh, it’s indeed negative, but you ignored smaller than or bigger than.

S4 laughed: Oh.

L: You did not answer question 5 and question 6. In question 7, P (Z = 3) = 0.9987, which table do you use?

S4: Type A.

L: S3, you are correct, probability of Z which *equals to* any value is always 0. Did teacher A mention this?

S3: Yes, he did, but he did not emphasise this because this situation is rare.

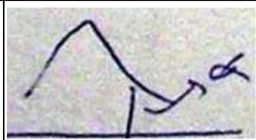
L: Let's test two questions of t distribution. Is this taught? How about chi-square distribution and F distribution?

S3: T distribution and chi-square distribution were taught last semester. But I don't remember F distribution being taught.

L: That's fine. We only look at t distribution today.

Q3: Look for t $_{(0.005,\ 9)}$, t $_{(0.99,\ 23)}$.

3 minutes later,

	t (0.005, 9)	t (0.99, 23)
Correct answer	3.25	-2.5
S3	3.250	
S4	3.250	1.319

L: S3, do you have an answer for t (0.99, 23)?

S3: No.

L: Both of your answers of t (0.005, 9) are correct. I know some students may look at 0.05 mistakenly. The table of t distribution is similar to normal distribution. There is no 0.99 in the table, and I saw you drew a graph S3.4, but the graph did not match the situation. α is the probability, and when it is 0.99, its area should take a large portion. So t (0.99, 23) should be negative, and the area of its left side is 0.01. Then you can find -2.5 in the table.

S4 laughed: Oh I looked at 0.1.

L: But your answer is positive.

S4: Yes. I subtracted 1 by 0.99

L: Why 0.1? Did you mistakenly write 1 – 0.99 = 0.1, or did you get 0.01 but looked at 0.1 in the table?

S4: I got 0.01 in calculation, and found the answer in the table.

L: And ignored it is negative.

S4: Yes.

L: That's fine. I know teacher A did not emphasise t distribution in his classes.

Q4. How did you realise the principles? What helped you in the process?

L: S3, I know that you draw graphs, and S4, you…

S4: I only use the table.

PART II

Q5. How do you define visualisation? What is included in your definition of visualisation?

S4: Drawing graphs.

S3: I think visualisation includes what you can see through your eyes, and it may not be a graph, but a formula, or a theorem.

L: So you mean what you can see?

S3: Yes.

L: Any more opinions?

S3 and S4: No

Q6. Did you use visualisation in your statistics learning? If yes, how and why? If no, why not?

L: S3, I can see you used it in the previous test. How about you, S4?

S4: No.

L: That's fine. S3, how do you use it, and when?

S3: For some formulas, no matter how hard you try to remember, you may still forget it after a long time. Thus, I wish to understand the formula by using the graphs, because I think *our brain can remember the graphs better*. I want to indirectly remember the formulas by remembering the graphs.

Q7. Did you use visualisation in the mid-term? If yes, does it help and how? If not, why not?

S4: No.

S3: Yes.

L: S3, you drew a different graph to that in teacher A's explanation of mid-term preparation. Why did you draw this?

S3 laughed: Because I don't rely on memorising formulas.

Q8. In what situations from the beginning of the course to the chapter of estimation, is visualisation needed the most needed by you?

S4: Standard variation and t distribution.

L: What do you mean by standard variation?

S4: I don't know.

L browsed textbook: Do you mean standard variation or standard normal distribution.

S4 pointed at standard normal distribution: This.

L: This is standard normal distribution.

S4 laughs.

L: Perhaps you don't remember its name correctly, but you didn't draw it.

S4: No.

L: You think visualisation is needed but you did not draw it, so you did not learn it well?

S4: I think so. I think visualisation is necessary here but I don't understand it.

S3: I think visualisation is more often used in fundamental mathematics.

L: What do you mean?

S3: For instance, the bar chart must be drawn. I also think that graphs must be drawn to deal with questions of probability.

L: What do you mean by probability?

S3: 'And' and 'or'.

Q9. Do you find/feel any function/advantage of visualisation that non-visual

methods cannot serve or cannot serve so well?

S4: A drawn graph makes the ideas clear.

L: Why don’t you learn how to draw?

S4: I don’t understand.

S3: It’s easier to remember the graphs and not easier to forget it.

PART III

Q10. What type of table and graph of standard normal distribution did you learn in the first semester? Did teacher C use type B as well as teacher A? Do you know other type(s)?

S3: Both type A and type B.

L: Which type did he use more often?

S3: Teacher C did not have a particular type.

L: What did he do? For example, how did he deal with P (Z > -a)?

S3: He would transfer to an easily dealt-with form.

L: I know that teacher A prefers type B, but do you know which type teacher C used more often?

S3 and S4: No

L: Ok, which type do you prefer?

S3: I always draw graphs of type B.

L: Do you understand type A?

S3: Yes, but normally I don’t use type A.

L: Did you use type A when solving questions before this interview?

S3: I would use type A in question 6.

L: But you drew a graph of type B.

S3: I don’t like > 0 and < 0.

L: What do you mean?

S3: I am not clear with the idea of bigger than and smaller than.

L: Even if with the help of graph?

S3: With a graph, I can have a better understanding, but I still need some time to think about it.

L: In question 6, you transferred smaller than to bigger than.

S3: There is another method. I can write smaller than this (-0.03) subtracts smaller than that (-2.41).

L: That’s correct, but you didn’t use this method.

S3: I used this method before, but I often made mistakes.

L: You mean you often made mistakes when you used smaller than this subtracts smaller than that?

S3laughed: Yes. Because I have to consider which one is bigger.

L: So you use teacher A’s principles to transfer to bigger than or 1 – bigger than?

S3: Yes.

L: You mean you have to think *whether a or b is bigger* in P (a < Z < b), and the bigger one subtracts the smaller one?

S3: Yes.

L: S4, which type do you prefer?

S4: Type A.

Q11. If you know more than one type, which one do you prefer and why? How well do you understand the types? Can you compare their strength and weakness?

S3: I can directly use the table of type A when facing negative or smaller than questions. I have to transfer bigger than negative to smaller than or 1 subtracts positive, then use type B.

Q12. Did you have difficulty in using the table? If yes, describe it.

S4: When it is not in the table.

L: You use type A which includes both positive and negative parts. Can you say what is not in the table?

S4: …

L: In question 2, P (Z < 1.22), 1.22 is in the table and smaller than is the same to type A. Do you have a reason why you did not answer this question?

S4: I was not sure if I was right, so I did not write it down.

PART IV

Q13. Could you please explain your understanding of the process of normalisation?

S4: Is it regrouping data?

L: Can you say again?

S4: It will become a positive and a negative.

L: How?

S4: ….

Q14. Could you please explain your understanding of the process of normalisation by graph?

S3 and S4: …

Q15. Do you remember the three z_{α} scores and their corresponding confidence intervals, do you know how to get them? Can you explain the process?

L: Do you remember these z scores?

S3: Teacher A asked us to remember them.

L: In the mid-term, did you remember these z scores or did find them in the table by yourself?

S4: I memorised them.

L: Can you find them in the table?

S3: Yes.

L: Ok. Can you find how much the z score is corresponding to 90% confidence interval?

S3 and S4 cannot find it in 1 minute.

L: Perhaps you can try 95% confidence interval.

S3: There is not 0.95 in type B, so I use 1-0.95. They are mutually matched.

[I think she means symmetry].

L: S4, are you pointing at 0.05 on the top of the table?

S4: The column below 0.05.

L: S3, what are you looking at?

S3: It is the middle between 0.0505 and 0.0495, so it is 0.05.

L: Right. So the *z score corresponding to 95%* confidence interval is?

S3: *1.645*.

L: How about 90% confidence interval? Just write it down.

S3: 0.1…

S3 cannot find 0.1 in the table.

S3: Where is 0.1?

L: Can you write down the three z scores that teacher A asked you to memorise?

S3: I think I can.

S3 wrote 90% 1.645, 95% 1.96, and 99% 2.575.

L: Do you notice that you wrote 1.645 for 95% previously? What is wrong?

S3 laughs.

L: S4, do you know what S3 wrote mean?

S4: That is different to what we memorised. If I am right, what S3 just wrote should be correct.

L: Right. I'll just briefly explain. 0.05 includes both sides, so each side is 0.025. When you looked for 0.05, it is only one side, and another side is also 0.05, so it corresponds to 90% confidence interval. Do you understand what I said?

S4: No.

S3: Yes

L: Ok, S3, you can tell S4 later.

S3 and S4 laugh.

Q16. Do you know what confidence interval means? What is the difference between 0.95 and 0.90 confidence intervals? Explain in your language.

S3: I think 0.95 confidence interval is wider than 0.90 confidence interval.

L: Any more opinions?

S3: No.

S4: No.

L: Ok, thank you so much. I am sorry for taking such a long time. I will shorten the time of later interviews, see you later.

S3 and S4: Bye.

[This interview took about 75 minutes. S3 and S4 had to leave for personal reasons.]

6.6. The first student interview with S5 on 1st May 2009.

PART I

Before the interview, S5 was given questions to answer:

Q1. Please look for probability in the table of normal distribution. (1) P (Z < -0.03), (2) P (Z < 1.22), (3) P (Z > -3.1), (4) P (Z > 2.05), (5) P (-1.4 < Z < 2.32), (6) P (-2.41 < Z < -0.03), and (7) P (Z = 3).

After 15 minutes, he finished.

L: Let’s check your answers. Question (1) and (2) are correct. Question (3), which table do you use?

S5: Type B.

L: Can you point at where the answer is in the table?

S5 pointed at 0.001, when z_α = 3.10: Here, 0.001.

L: Oh, it should be -3.1, your table use is correct.

S5: I am sorry.

L: That’s fine. Question (4), you looked at the neighbour of the correct answer.

S5 laughed: Ha hah, I was not careful enough.

L: Question (5), P (Z > 2.32) = 0.0102, and question (6), P (Z > 2.41) = 0.008. You looked at their neighbours again. In the first six questions, you looked at the neighbour of the correct answer four times.

[In Q1.1, S5 also found the neighbour of correct answer, but he corrected it when he noticed it.]

	P(Z<-0.03)	P(Z<1.22)	P(Z>-3.1)	P(Z>2.05)	P(-1.4<Z<2.32)	P(-2.41<Z<-0.03)	P(Z=3)
Correct Answer	0.488	0.8888	0.999	0.0202	0.9898 - 0.0808 (or 1 –0.0808 –0.0102) = 0.909	0.488 – 0.008 [or (1 - 0.008) - (1 - 0.488)] = 0.48	0
S5 Type B	P(Z<-0.03) = ~~0.4920~~ = 0.4880	P(Z<1.22) = 0.8888	P(Z> *3.1*) = *0.0010*	P(Z>2.05) = *0.0207*	P (-1.4 < Z < 2.32) =P(Z>-1.4)-P(Z>2.32) =1-P(Z>1.4)-P(Z>2.32) =1 - 0.0808 - *0.0104*	P(Z>-2.41)-P(Z>-0.03) =1-P(Z>2.41)-1-P(Z>0.03) =1 - *0.0082* – 1 - 0.4880	P(Z=3) = *0.0013*

L: Can you use the table of t distribution?

S5: Yes, I can.

L: Ok, please answer the two questions.

Q3. Look for t $_{(0.005, 9)}$, t $_{(0.99, 23)}$.

[3 minutes later.]

L: The first one is correct; you know something about the second one, but the answer is only -2.5. The 0.99 is not in the table; 0.99 is the area in the right-hand side of the particular value, and the other area (left-hand side of the value) is 0.01, you can find

0.01 in the table and you just need a negative symbol.

	t (0.005, 9)	t (0.99, 23)
Correct answer	3.25	-2.5
S5	t (0.005, 9) = 3.250	t (0.99, 23) = 1 – 0.01 = 1 – 2.500

Q2 & Q4. How did you realise the principles? What helped you in the process?

S5: Do you mean how I calculate them?

L: Yes, such as why you write 1 – something.

S5: There is no negative z score in the table of normal distribution (TYPE B), so according to teacher A, you should transfer the negative one to a positive one, and the first step of P (Z > -a) is 1 – P (Z < -a) = …. I don't know how to explain.

L: That's fine. Can you do it by yourself without your notes or textbook?

S5: Yes.

L: Do you have an image of it?

S5: Yes, I already have an image. The teacher gave us the formulas. He did not explain where the formulas come from but he asked us to remember and use them.

L: In the first semester?

S5: No, in this semester.

L: Were you taught this part in the first semester?

S5: No.

L: How about using the table?

S5: Using the table was taught in the first semester, but calculation was not.

L: So you were only taught how to use the table, but how to get anything not in the table was not taught?

S5: No.

L: What did you use to help solve the question in Q1? I mean anything other than the principles of calculation, and the table of normal distribution?

S5: I use the formulas given by teacher A.

L: Ok, these are about Z distribution, let's talk about t distribution now. There are also some principles in t distribution. How do you deal with the values that are not in the table, such as 0.99? Can you explain why you write "1 –" in the last question?

S5: I found 0.01 but not 0.99, so 1 – 0.01 = 0.99, and 0.01 is corresponding to 2.5

L: Do the different representations of Z and t distribution cause your difficulty?

S5: Yes, and I think that 1 – 0.99 = 0.01

L: Ok, thanks. The first part is finished.

PART II

L: In the first part, when you solve the questions in Q (1) and Q (3), you did not draw any graph.

S5: There are graphs drawn on top of the tables, so I directly use them.

L: But the graph only represents the “bigger than” and “positive” part; and for t distribution, 0.99 is neither in the table, nor in the graph. It is not wrong that you did not draw a graph; I mention it because it is in my research area.

S5: I know

Q5. How do you define visualisation? What is included in your definition of visualisation?

S5: For example?

L: Such as the graph and table.

S5: Both teachers taught us how to draw graphs, but formulas were more often given and used. How teachers solve a question is then used by us to solve the *next* question without thinking of using or drawing graphs. We *imitate* methods used in the previous question without really understanding.

L: I see.

Q6. Did you use visualisation in your statistics learning? If yes, how and why? If no, why not?

S5: In fact, I also know how to draw the graphs, but…

L: But you did not draw.

S5: No. hah…

Q7. Did you use visualisation in the mid-term? If yes, does it help and how? If not, why not?

S5: Yeah. Wait, let me see… no, I didn’t.

L: I see. May I ask why?

S5: Because the questions can be solved by calculation, and only the answers are different.

L: Do you know that 80 percent of questions are the same as the preparation paper?

S5: Teacher A said that some questions will be the same, but he didn’t say how many of them.

Q8. In what situations from the beginning of the course to the chapter of estimation, is visualisation needed the most needed by you?

S5 browses his textbook and notes.

S5: I think it becomes important in Chapter four (probability). There are many events including empty event, intersection event, mutual exclusive event [i.e. set], and drawing graphs can help explain the ideas. Besides, conditional probability and independent events are more difficult ideas and the teacher drew graphs to help us understand. I feel that graphs are necessary from chapter four to the present content [i.e. confidence intervals].

Q9. Do you find/feel any function/advantage of visualisation that non-visual methods cannot serve or cannot serve so well?

S5 browsed his book for 3 minutes.

S5: One of the disadvantages of visualisation is that I may be unable to understand the graphs drawn by myself.

L: You may draw some graphs that you don't understand?

S5: Yes. Perhaps the teacher has taught, but I don't know how to draw a graph for a *new* question.

L: Any advantage?

S5: Such as the attached practice question in page 312, graphs help us to know how to solve it, but sometimes graphs are difficult to understand.

PART III

Q10. What type of table and graph of standard normal distribution did you learn in the first semester? Did teacher C use type B as well as teacher A? Do you know other type(s)?

S5: I use type B.

L: What type did teacher C use?

S5: He taught us both type A and type B.

L: How about teacher A?

S5: Both.

L: Which type is more often used?

S5: I think… type B.

L: Which teacher do you mean?

S5: Both of them.

L: I see. Do you know any other type?

S5: I don't think so.

Q11. If you know more than one type, which one do you prefer and why? How well do you understand the types? Can you compare their strength and weakness?

L: As you just said that you know two types, which type do you prefer?

S5: I prefer type B.

L: Why?

S5: Because there is *no negative* symbol in the table, and it *looks easier to calculate*.

L: Is it easier to calculate without a negative symbol?

S5: Yah, it's just my feeling.

L: What if there is a negative symbol in the question?

S5: Hmm… then I have to transfer. It's just my personal feeling. I feel it is easier to calculate without the negative symbol.

L: That's ok. Can you list the advantage and disadvantage of both types?

S5: I think the advantage of the tables is that you can find answers in the tables.

L: Disadvantage?

S5: I think people in the present day have difficulty in interpreting or making this table, they may feel that it is troublesome.

L: Can you compare type A and type B?

S5: We can use type B correctly, but many people may not notice the existence of the second page of type A. Many of them may not understand where the table comes from, and they may just use it. Without the table, we can still use other methods to find the answers. Many people don't use the table, and that is the disadvantage.

L: When you say table, do you mean graph or table?

S5: Oh, I mean graph.

L: And you can find the answers by using the formulas?

S5: That's right. I feel that the graph is not necessary especially when people know how to calculate with the formulas.

Q12. Did you have difficulty in using the table? If yes, describe it.

L: I just notice that you often look at the value next to the correct answer. Do you have any other difficulty when using the table?

S5: If I know how to use the table and the steps, then there is no other problem.

L: Ok, thanks.

PART IV

Q13. Could you please explain your understanding of the process of normalisation?

S5: I don't know what the word normalisation means.

L: I know that you were taught and used the formula Z = (X – μ) / σ and Z = (X bar – μ) / (σ/√n), do you know what the formulas mean?

S5: For me, normalisation (standardisation) offers a standard. For example, teachers should give a standard for examination, such as 60 or 70 are required to pass.

L: If X ~ N (200, 20) and you normalise X, what is the expected value of Z?

S5: Do you mean the mean of (X – 200) / 20?

L: Yes. What is the mean of standard normal random variable Z?

S5: I don't know what this means.
L: That's fine. What is the standard deviation of Z?
S5: Isn't it 20?
L: For standard deviation of Z?
S5: I really don't know.
L: Don't worry.

Q14. Could you please explain your understanding of the process of normalisation by graph?

S5: You mean Z = (X – 200) / 20?
L: Yes.
2 minutes later.
S5: I am not sure but I would like to try.
L: Go ahead.
[S5 drew a graph: a curve on top of a straight line. he marked the line and wrote 0 below the mark, then marked and wrote 20 in the right-and side of 0. Finally he marked and wrote x-200 in the right-hand side of 20. He looked confused.]
2 minutes later.
L: This question may take you quite a long time.
S5: Yes, and we rarely notice how the graphs are drawn, and only directly use the formula to calculate, but I would like to try drawing.
L: Ok, thanks.

Q15. Do you remember the three z_α score and their corresponding confidence intervals, do you know how to get them? Can you explain the process?

L: Do you remember the three z_α scores or do you find them in the table?
S5: I memorise them, but I also know how to find them in the table.
L: Ok, can you find the z_α scores corresponding to 95% confidence interval in the table?
S5 wrote 95% = 1.96
S5: 95%, I should look for 0.025 in the table.
L: What is 0.025?
S5: 0.025 is that in the table, and I can find it between 1.95 and 1.96; 0.0256 corresponds to 1.95 and 0.025 corresponds to 1.96, so the answer is 1.96.
L: How about 90% confidence interval?
S5 wrote 1 – 0.9 = 0.1 and 0.1/2 = 0.05, then he spent 2 minutes to look for the answer in the table.
S5: I think it is between 1.945, right?
L: Can you write it down?
S5: Ok, 90% confidence interval…, so 1 – 0.9 = 0.1, 0.1 / 2 = 0.05, and then look for

0.05 in the table. As you can see, 0.05 is between the two values corresponding to 1.94 and 1.95,
L: Can you show me in the table?
S5: Here.
L: Why do you point at 0.0256?
S5: 90% confidence interval corresponds to 1.96, so I think 95% will be nearby.
L: Since you can find 1.96 corresponding to 0.025, why don't you just look for 0.05 in the table?
S5: Because 0.025 corresponds to 1.96, so I think 0.5 will be *close* to it.
L: Do you mean that 0.05 and 0.025 are very close?
S5: Yes.
L: Can you try again to look for 0.05?
S5 pointed at .05 on top of the table, not probability in the table: This?
L: Not the outside one, it's inside.
S5: It takes a long time, because I have to search one by one.
[S5 did not notice the increasing tendency of the probabilities.]
S5 pointed at .0548, .0537, .0526, .0516, and .0505: These are all 0.05.
L: 0.05?
S5: Do you mean 0.050000?
L: Yes, can you find it?
S5: I don't know, and it takes a long time….
S5: I can find 0.0005, but I cannot find 0.05…
L: Which one is closest to 0.05?
S5: All in this row?
L: And which one is really closest to 0.05?
S5: 0.0505?
L: Anything else?
S5: I don't know.
L: Can you write it down?
S5: Ok.
S5 wrote 0.0505 = 1.64
L: That's great. Can you write down the three z_α scores in your memory?
S5 wrote 95% = 1.96; 90% = 1.945 and 99% = 2.575
L: Thank you so much.
S5: Any problem?
L: Your memory may not be so reliable, as you can see that 90% should be 1.645.
S5: Oh, right.

Q16. Do you know what confidence interval means? What is the difference

between 0.95 and 0.90 confidence intervals? Explain in your language.

S5: It's relevant to the intervals of 90% and 95%. When you calculate 1 – 0.95, its confidence interval is located at 0.05.

L: You mean 1 – 0.95 = 0.05, and it is the confidence interval?

S5: Yes, its confidence interval is at 0.05.

L: That's ok, what is the difference between 0.95 and 0.90 confidence intervals?

S5: The numbers are different and the results will be different. The length of intervals will also be different.

L: Which one is longer?

S5: 95%?

L: Why?

S5: Because 90% is 1.645 and 95% is 1.96, and 1.96 is longer than 1.645, so 95% is longer.

L: Ok, thanks. This interview ends here. See you next Tuesday.

S5: See you.

6.7. The first student interview with S6-S10 on 4th May 2009.

PART I

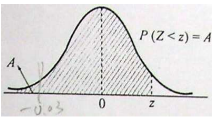
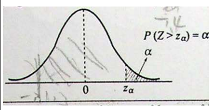
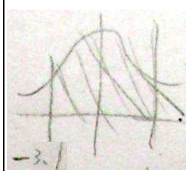
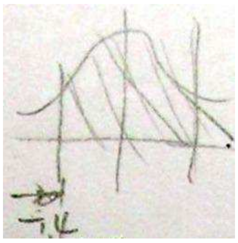
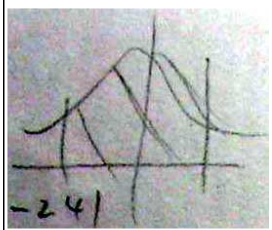
Before the interview, S6 - S10 were given questions to answer:

Q1. Please look for probability in the table of normal distribution. P (Z < -0.03), P (Z < 1.22), P (Z > -3.1), P (Z > 2.05), P (-1.4 < Z < 2.32), P (-2.41 < Z < -0.03), and P (Z = 3).

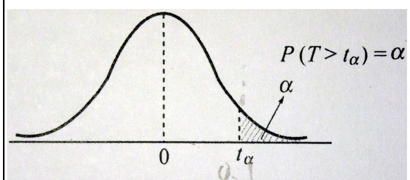
Q3: Look for t $_{(0.005, 9)}$, t $_{(0.99, 23)}$.

After 20 minutes, all of them finished answering.

L: Let's check your answers.

	P(Z<-0.03)	P(Z<1.22)	P(Z>-3.1)	P(Z>2.05)	P(-1.4<Z<2.32)	P(-2.41<Z<-0.03)	P(Z=3)
Correct Answer	**0.488**	**0.8888**	**0.999**	**0.0202**	**0.9898 - 0.0808** **(or 1–0.0808–0.0102)** **= 0.909**	**0.488 – 0.008** **[or(1 - 0.008) - (1 - 0.488)]** **= 0.48**	**0**
S6 Type B	P(Z<-0.03) =P(Z>0.03) =0.488   [S6 directly marked -0.03 on the graphs in the textbook]	P(Z<1.22) = 0.8888	P(Z>-3.1) =1-P(Z>3.1) =1-0.001 =0.999  	P(Z>2.05) =0.0202	P (-1.4 < Z < 2.32) =P(Z>-1.4)-P(Z>2.32) =1-P(Z>1.4)-P(Z>2.32) =1-0.0808-0.0102 =0.909  [This graph is the same to the graph in Q1.3, S6 just replace -3.1 by -1.4.]	P(-2.41 < Z < -0.03) =P(Z>-2.41)-P(Z>-0.03) =1-P(Z>2.41)-1+P(Z>0.03) =1-0.008-1+0.488 =0.48  	P(Z=3) =1- (Z<3) -P(Z>3) =0
S7 Type B	P(Z<-0.03) =P(Z>0.03) =0.488	P(Z<1.22) =1-P(Z>1.22) =1-0.1112 = 0.8888	P(Z>-3.1) = P(Z>-3.1) =1-P(Z>3.1) =1-0.0010 =0.999	P(Z>2.05) = P(Z>2.05) =0.0202	P (-1.4 < Z < 2.32) =P(Z<2.32)-P(Z>-1.4) =1-P(Z>2.32)-P(Z>1.4) =1-0.0102-0.0808 =0.909	P(-2.41 < Z < -0.03) =P(Z<-0.03)-P(Z>-2.4) ~~=1-P(Z>0.03)-P(Z>2.4)~~ ~~= -0.488-0.0082~~ ~~= -0.5038~~ =>P(Z>0.03)-P(Z>2.4) =0.488-0.0082 =0.4798	P(Z=3) =0?

S8 Type B	P(Z<-0.03) =P(Z>0.03) =0.4880	P(Z<1.22) =P(Z>1.22) =0.1112	P(Z>-3.1) =P(Z>3.1) =0.9990	P(Z>2.05) =0.0202	P (-1.4 < Z < 2.32) =	P(-2.41 < Z < -0.03) =	P(Z=3) = 1
S9 TypeB	P(Z<-0.03) =P(x-μ/σ<-0.03) =P(x<-0.03) = ~~1 –~~ P (x > 0.03) = ~~1 –~~ 0.4880 ~~= 0.5120~~ [The initial answer was 0.5120. S9 deleted “1-” in last third and second columns and left 0.4880 as the answer.]	P (Z < 1.22) = P($\frac{x-\mu}{\sigma_x}$ < 1.22) = P (x > 1.22) = 0.1112	P (Z > -3.1) = P($\frac{x-\mu}{\sigma_x}$ > -3.1) = P (x > -3.1) = 1 – P (x > 3.1) = 1 - 0.0010 = 0.9989	P (Z > 2.05) = P (x-μ/σ > 2.05) = P (x > 2.05) = 0.0202	P (-1.4 < Z < 2.32) =P(x>-1.4)-P(x>2.32) =1-P(x>1.4)-P(x>2.32) =1-0.0808-0.0102 =0.9192-0.0102 =0.909	P(-2.41 < Z < -0.03) =P(x>-2.41)-P(x>-0.03) =1-P(x>2.41)-1-P(x>0.03) =(1-0.008)-(1-0.488) =0.992-0.512 =0.48	P(Z=3) =0.0013
S10 Type B & negative part of type A	P(Z<-0.03) =0.488	P(Z<1.22) = 1-P(Z>1.22) =1-0.1112 =0.8888	P(Z>-3.1) =1-P(Z>3.1) =1-0.001 =0.999	P(Z>2.05) =0.0202	P (-1.4 < Z < 2.32) =1-(P>2.32)-(Z<-1.4) =1-0.0102-0.0808 =0.909	P(-2.41 < Z < -0.03) =P(Z>0.03)-(Z<-2.41) =0.488-0.0207 =0.4673	P(Z=3) =1

	t (0.005, 9)	t (0.99, 23)
Correct answer	**3.25**	**-2.5**
S6	t (0.005, 9) = 3.25	t (0.99, 23) = 31.821 – 2.5 = 29.321 [S6: I guess.] 
S7	t (0.005, 9) = 3.355 [S5 thought v = 9 - 1 = 8.]	t (0.99, 23) = -2.508 [S5 thought v = 23 – 1 = 22.]
S8	t (0.005, 9) = 3.250	t (0.99, 23) = -1.500 1-0.99=0.01 1-2.500=-1.500
S9	t (0.005, 9)	t (0.99, 23)

S10	t (0.005, 9) = 3.25	t (0.99, 23) =31.821-2.5-2.5 =26.821 [S10: I guess.]

Q2 & Q4. How did you realise the principles? What helped you in the process?

L: I wish each of you to respond and speak in the order from S6, S7… to S10.

[In their second interview, I changed the order because I found that they would copy the previous interviewees' words and ideas. Thus, I let students who did not write much speak first, and let students who wrote more speak later. By doing this, I think I obtained more original thoughts from them.]

S6-10: Ok.

L: S6, please.

S6: I drew graphs.

L: But you did not draw.

S6: Yes, I did.

L: Where?

S6 pointed at his three graphs drawn on the blank area above the tables: Here.

L: Then?

S6: I drew graphs to realise how to get the answer.

L: You are the first one who has said that. I know that you also drew graphs in the mid-term exam paper. Ok, S7, please.

S7: I decide which table to use according to whether the question asks > or <, but in most questions, I use type B (>) taught by teacher A. If the question asks <, I will use "1-" to become >.

L: Which table do you use?

S7: This (type B).

L: Do you also use type A?

S7: I would use it to check my answer to the question involving the negative symbol.

L: Only to check?

S7: Yes. I use type A to check answers of question asking "smaller than a negative one in that side".

L: I see, S8, please.

S8: I make a decision according to "bigger than" and "smaller than".

L: Which table do you use?

S8: This one (type B). If there is a "negative value", I will write "1 – positive value".

L: What if the question asks "smaller than"?

S8: That is what I don't know.

L: Ok, S9, please.

S9: I also use the formulas and the table.

L: Is there any question that you find difficult?

S9: …

L: Which table do you use?

S9: This (i.e. type B).

L: Do you use type A?

S9: No.

L: Ok, S10, please.

S10: I also make a decision according to > or <.

L: Which type of table do you use?

S10: I use the 2 pages (type B and the negative part of type A).

L: Don't you use the positive part of type A?

S10: No.

L: How about your principles of calculation?

S10: The same as them.

L: Ok. Let's talk about t distribution. Is there any difference between table of t-distribution and z-distribution, S6?

S6: I only feel that table of z distribution is easier to search.

L: why?

S6: Perhaps I have more impression of it so I can easily find answers with the table, but I don't understand table of t-distribution well. For instance, I don't know how to deal with the second question of t distribution.

L: Do you know where 0.99 is in the table?

S6: No, I don't know.

L: So what should you do?

S8 interrupted: You only can rely on imagination.

L: S6, did you draw any graph in Q (9)?

S6: Yes, but I still don't understand.

L: Where did you draw?

S6 pointed at the graph attached on the top of the table in his textbook: Here. I thought I can get the answer from 1 minus the answer of 0.01.

[Some sentences here and below are not easy to understand, because students frequently said this, that, or other nouns and un-structured sentences.]

L: Did you answer it?

S8: Yes, there is a negative one in front of it.

S10: It becomes negative.

L: Are you sure about your answer now?

S6: No.

L: That's fine, S7?

S7: I feel that t-distribution is easier. But I have to consider v which I don't know how to deal with. Normally teacher A gives a. zero point something followed by a number, and it is easier to find the answer than z distribution which requires drawing graphs.

L: Did you draw any graph?

S7: No.

L: So you remember the principles?

S7: Yes, and judge.

L: Ok, what's your answer to Q (9)?

S7: The other one of 0.99 is 0.01, so I directly looked at 0.01 in the table and then added a negative symbol.

L: So, what's your answer?

S7 laughed: -2.508

L: Is the v that you look at 22?

S7: Yes, I didn't know that I should directly look at the v given in the question. I remembered that there is a "- 1" to get v.

L: Did teacher A mention that?

S7: Yes, but I don't remember exactly.

L: That's fine, you should just look at v = 23.

S8: Oh?

S7: Really?

S10: Yeah.

S6: 1 - ?

L: How about you, S8?

S8: I don't feel that they are too different. Basically if the number is exactly on the point, I can find it. When there is a negative symbol, I will use 1 – the number corresponding to the positive one in the table. But for this question Q (9), 1 – 0.99 = 0.01, then 1 – the number corresponding to 0.01 would be the answer of 0.99.

L: Don't you have the process? Or do you use mental calculation?

S8: I calculate on the side.

L: I didn't see it, did you erase it?

S8 laughed: Yes.

L: Oh, no! I told you not to erase anything.

S8 laughed: I forgot.

L: What's your answer?

S8: -1.5.

L: Can you write down again how you get it?

S8: Ok, I write it here.

L: Good, thanks. S9, it's your turn.

S9: I cannot solve Q(8) and Q(9).

L: Both of them?

S9: Yes.

L: That’s ok, S10?

S10: I only use the table directly.

L: Do you have answers?

S10: I had an answer for Q (8), but I didn’t write down the answer for Q (9) because I only guessed it.

L: Please write it down and note that you guessed it.

S10: Ok.

L: Thanks, we just finished the first part.

PART II

Q5. How do you define visualisation? What is included in your definition of visualisation?

S8: I realise that when I see it.

S6 asked S8: Do you realise it or hate it when you see it?

L: Good ideas, I will write it down. Do you have any explanation?

S6: When I see some graphs, I know that they are very difficult right away, so I hate graphs. But I am still used to drawing graphs in the beginning when solving questions.

L: You mean you hate Q (9)?

S6: Yes, I feel that this question would be very difficult, but I still tend to draw graphs.

L: Do you know the meaning of the graphs?

S6: I may spend much time thinking what the graph is about.

L: Ok, thanks, S7?

S7: In statistics, I like some graphs and also hate some other graphs.

L: What graphs do you like and what graphs do you hate?

S7: I prefer graphs in previous chapters, such as histograms. I really hate graphs of standard normal distribution that require judgment. Graphs of intersection and union of sets are necessary for thinking. I don’t like them but I still need to draw them to help understanding. However, I don’t like the graphs of standard normal distribution, and I also don’t want to learn it.

L: Why?

S7: You may use graphs to solve questions, but you can also solve questions without drawing graphs. So I feel that I don’t have to learn it.

L: I see. S8?

S8: I feel that some graphs let me have a better understanding of what is being talked about than words do.

L: Such as?

S8: Such as intersection and union of set. The drawn graph lets me better know what is called intersection or union than only words.

L: How about later chapters?

S8: Graphs in later chapters are more difficult. Perhaps it is also influenced by my interpretation of the teacher's speech in class.

L: Do you use graphs in normal distribution?

S8: No, I don't, but I feel that graphs are very important in the chapter on intersection and union.

L: Thanks, S9?

S9: I don't know graphs because I didn't memorise how to draw and interpret graphs. I only remember how to solve by using formulas.

L: How do you remember the formulas?

S9: I memorise them when I write them.

L: If the question asks "> negative"?

S9: I use "1 -"

L: What if it asks "<"?

S9: … "Z <" equals "Z >"…, it may be wrong…

L: That's fine. S10, please.

S10: Basically, if I see a graph, I will just ignore it. If it is necessary, I will remember it in my own method.

L: What do you mean your own method?

S10: I don't know how to describe it.

L: Which question that can be solved with graphs can be solved by other methods without graph?

S10: I don't use graphs. I solve each question by my own method.

L: Ok, we will see how your own method works later.

Q6. Did you use visualisation in your statistics learning? If yes, how and why? If no, why not?

L: Some of you already talked about this issue before I asked. S6, you say that you hate graphs but you still learn it and use it in examination.

S8 laughed: I didn't hate it but I didn't learn it.

L: Are you serious?

S7: It depends on which chapter it is.

S6: Yes, it depends on which chapter it is.

S7: Also depends on my mood.

L: Ok, next question.

Q7. Did you use visualisation in mid-term? If yes, does it help and how? If not, why not?

L: I know that there are graphs on S6's mid-term paper.

S6: Yes, I draw if it is necessary, but I don't draw for common calculation. Graphs are necessary in many questions, and I will understand what the questions are about by my own.

L: I know that most questions in the mid-term were the same as questions in the preparation paper, and many students memorised the answers, I wonder if you memorised the answers and graphs or realised how to solve them?

S6: I didn't memorise the answers because I may lose some processes in my memory.

L: Ok, thanks. Do you have any more ideas?

S6-10: …

L: Ok, next question.

Q8. In what situations from the beginning of the course to the chapter of estimation, is visualisation needed the most by you?

L: Some of you also mentioned this issue already. Do you want to add any comment, S6?

S6: I have no comment to add.

L: In what content you feel graphs are the most necessary?

S6: Many of them. If the question mentions it, then I need it.

L: Ok, thanks. S7?

S7: There are two kinds of graphs: bar-chart, pie-chart in earlier chapters, and those which require calculation and making decisions in later chapters.

L: Graphs in the earlier part are presenting data.

S7: I like to use graphs in earlier chapters, but in later parts, graphs are more difficult and I don't understand so I just let it be.

L: Thanks. S8?

S8: No comment.

L: S9?

S9: No.

L: That's fine, S10.

S10: I don't use graphs.

L: May I ask why?

S10: Because drawing graph is troublesome.

L: But some graphs are easy.

S10: I just don't want to.

L: Thanks.

Q9. Do you find/feel any function/advantage of visualisation that non-visual methods cannot serve or cannot serve so well?

S6: Sin and cos. Its idea cannot be understood without graph.

S8: The way of presenting data such as on page 43 of the textbook helps me to understand more easily. Taking the collected data for example, I can easily see which is/are more or fewer in a peek if it is presented in a bar-chart. It's the advantage of the graph. Without the graph, you have to count each of them, but with the graph, you only need a peek.

L: Thanks, S9?

S9: No comment.

L: S10?

S10: No.

L: Ok, thanks, the part II is over.

PART III

Q10. What type of table and graph of standard normal distribution did you learn in the first semester? Did teacher C use type B as well as teacher A? Do you know other type(s)?

S7: Teacher C taught both types.

S6: Yes, he taught both.

S7: He explained that they are table of "bigger than" and of "smaller than" respectively, but type B is more frequently used.

L: Do you know any other type?

S6-10: ….

Q11. If you know more than one type, which one do you prefer and why? How well do you understand the types? Can you compare their strength and weakness?

S6: I prefer type B, but if the answer of asking "smaller than" can be directly found in type A, I will use it without troublesome transferring and calculation.

L: Thanks, what are the advantage and disadvantage of type A and type B?

S6: I don't know. I just feel that the important thing is being convenient to use.

L: Convenient to use, good idea. S7, please.

S7: The advantage of type B is that most questions are in this form.

L: You mean in teacher A's questions?

S7: Yes.

L: But "smaller than", "bigger than", "positive" and "negative" are all included in my questions.

S7: Type B is still easier to use. Perhaps I am used to type B, so if the question asks "smaller than", we will transfer it to "1 – bigger than". We don't use type A often, and the numbers are all negative, so it is more difficult to judge.

L: Don’t you know that numbers in the other page of type A are all positive?

S7: Well, I am not familiar with this table.

L: That’s ok. S8?

S8: In fact, in the two tables, it is easy to find if it is on the point.

L: Which type do you prefer?

S8: No difference, I don’t like or hate any type.

L: Which one is easier to find?

S8: Perhaps the negative one is easier to find.

L: Type A?

S8: Yes. It includes both negative and positive, so the questions should be more likely included and the answers can be directly found in the table.

L: Ok, thank you. S9?

S9: I prefer “bigger than”.

L: You mean type B. Because of its advantage, or other reasons?

S9: Because I use it more often in practice. I don’t bother to turn the page, so I only use this. When the question asks negative, I transfer it to “1 –”.

L: Ok. S10?

S10: Whatever.

L: When do you use type A and when do you use type B?

S10: I use whatever table I get.

L: No difference?

S10: No

L: Thanks.

Q12. Did you have difficulty in using the table? If yes, describe it.

S6: I only feel that too many numbers in the table make it difficult for me to find.

L: Too many numbers? In which type?

S6: In both types.

L: Ok, S7?

S7: When I use type B, I have difficulty if the question asked “smaller than” or “negative”. I have to decide whether I should use “1-” or transfer “negative” to “positive”.

L: Ok, S8?

S8: It’s difficult to find if it is not on the point. I have a problem when it involves transferring.

L: I see, S9?

S9: I may find a wrong one.

L: Like its neighbour?

S9: Yes.

L: Oh, any problem in transferring, like “1-”?

S9: No.

L: Ok, S10?

S10: The only problem is that the words are too small.

L: Do you mean that you may find a wrong one, or find a correct one but write a wrong one?

S10: Both.

L: I see, your problem is that the words are too small. We just finished part III.

PART IV

Q13. Could you please explain your understanding of the process of normalisation?

S6-10: …

L: For example, what is the purpose of normalisation?

S6: It will be easy to use the table.

S7: I don’t know, I learn it because it is taught.

S8: I don’t know its meaning, either. I only know that the answer can be gotten by calculation.

S9: I also memorise the formula, I know that $Z = (X – \mu) / \sigma$.

L: Do you know how this formula is generated?

S9: No.

L: How about you, S10?

S10: The formula is taught like this, so I write it like this.

Q14. Could you please explain your understanding of the process of normalisation by graph?

S6: No.

L: Others?

S7-10: No.

Q15. Do you remember the three z_α score and their corresponding confidence intervals, do you know how to get them? Can you explain the process?

L: Please prepare the table of normal distribution which you normally use. I know that teacher A asked you to remember the three z_α scores and their corresponding confidence intervals, can you find the z_α score corresponding to 95% confidence interval by yourself? Please “use the table”.

15 seconds later, S10 pointed at .0250 (-1.9 on left and .06 on top) in the first page of table of type A.

L: Ok, please write the z_α score on your paper.
40 seconds later, S9 pointed at 0.025 in the table of type B.
S8: Why cannot I find it?
S9 and S10 were laughing…
S7: Why is there such a question? I can memorise them, 90% is 1.575...
L: Stop, stop.
S8: 1.96.
S7: 2.5.
S6: 2.44, I also memorise them.
L: Do you memorise them?
S9: By calculation.
L: How about you, S10?
S10: Both calculation and memorising.

L: Can you find it?
S7: No, I cannot. Why is there such a question?
L: I know that teacher A asked you to memorise them, but I want to know how you find them.

S6, S7 and S8 spent more than 3 minutes but they were confused; S6 looked at the negative part of type A; S7 and S8 looked at the table of type B and the positive part of type A.
[Perhaps they also looked at other pages, but I didn't see or record.]

L: Ok, can you try drawing what 95% confidence interval means in the graph?
S7: From this [a vertical line drawn in the left-hand side of table B] to that [the vertical line printed in the right-hand side].
S6 drew a new graph with two vertical lines and pointed at the middle area: here.
L: Can you draw shadow to represent 95% confidence interval?

L: If the middle area is 95%, how much is other areas?
S6 wrote 0.05/2 = 0.025.
L: Ok, I think you get it.
S7: How do you find it?
S10: You found it?
S6 turned to look for 0.025 in the table of type B: not yet, not yet.
S7 was murmuring: 1.96…
L pointed at the probabilities near .0197: Are you looking for 1.96 in the table?
S7: There is no way to do it!

L: So where should you look?
S7: Look for two close numbers.
L: Where do you look now?
S7's pen pointed at .1977: here!
L: But that's 0.1977.
S7: Oh, no. I should look at… That's too difficult.
L: How about you, S8?
S8: I cannot find anything about that.

S9 and S10 were laughing and S6, S7 and S8 were blaming them not helping them.
L: Are you laughing because you find that some classmates don't realise this? I just noticed that this may be an obstacle for many students because other participants had difficulty, too.
S7: This is really hard.
S9: I also did not know in the beginning.
L: S6 has the answer now.
S6 pointed at .0250: this.
L: Right. What's the z_α score you were asked to memorise?
S6: …
L: You may not remember it now after using the table.
S6: 1.96.
L: Do you know what it means now?
S6: Yes.
L: Really understand?
S6 laughed: Really, it's so impressive!
L: Ok, I will write it down.
S8: Too difficult.
S10: You should ask them the one corresponding to 90% confidence interval. They may look for it till death.
L: Let me see your answer.
S10 pointed at 0.505 and 0.495: Between two of them.
L: You are right.
S8: Can you let me know the answer to the previous question?
S7: Can you also tell me?

S6, S7, S8 and S10 discussed the question.

L: Can you find the z_α score corresponding to 99% confidence interval?
S9 wrote 0.01/2 = 0.005, and pointed at .0005 initially, but his finger moved way and

then stopped at .0051 and .0049.

L: The answer is?

S9: 2.578.

L: Why 2.578?

S9 wrote 0.05 between .0051 and .0049: 0.05 is here. The left side is 2.5 and top…78.

L: 2.578.

S9: No, no, no. I was wrong. It's 2.575.

L: Good, thanks. I will also note your previous mistake.

When I asked S9 the z_α score corresponding to 99% confidence interval, S6, S7, S8 and S10 discussed as follows:

S7: I know it is 1.96, but what I find is this. [i.e. .0250]

S10: 0.05 should be divided by 2.

S6: Yes.

S10: And it becomes 0.025.

S7 and S8 asked at the same time and their voices are mixed, I cannot understand what they asked but according to S10's answer, they seemed to ask why.

S10: It has the possibility of positive and the possibility of negative, so divided by 2.

S8: So divided by 2… Then 90% confidence interval, it becomes 0.10, right?

S6 and S10: 0.1 divided by 2.

S10: Get 0.05. You may spend much time to find 0.05.

S7: Did S10 just say which one I cannot find, 90% or 99%?

S10: 90%.

S7: How about 99%?

S6: 0.01 divided by 2.

S10: 0.025.

S7: 0.025.

S8: It should be 0.005.

S6: Right, 0.005.

S7: 0.002?

S10: 0.005!

[The dialogue in this paragraph happened when I was asking S9.]

S8: Ok, for 90% you said 0.05.

S10: Can you find it?

L: This is an interesting question.

S10: More interesting than the previous one.

L: I asked the previous one because it is easier to find.

S8 moved her pen to point at the probabilities which are not getting close to 0.05.

Initially her pen stopped at .0051 and .0049, but soon moved away.

L: What are you looking for?

S8: 0.05.

L: Where is it?

S8: I am looking for it.

S7: Why is there such a 1.575?

L: Which one?

S7 pointed at .0250: I find 1.96 in this way, but why is there 1.575?

L: Do you have the answer?

S6 pointed at .0505 and .4995: 1.645.

L: That's of 90% confidence interval. How about 99%?

S6 pointed at .0051 and .0049: It's 2.575

L: Right.

S10: You get smarter.

S6 laughed.

L: Ok, we only have 5 minutes for the last question. Did you find them yet?

S7: I am not familiar with them.

S8: I am not sure.

L: Not sure? You can try. Which one do you think?

S8 pointed at .0052 and .0051: 0.050

L: So what's your answer?

S8 pointed at .0051: this one.

L: Can you write it down?

S8 wrote 99 between .0051 and .0049.

L: What's the z_α score which you were asked to remember?

S7 pointed at .0505 and .4995: Is the answer of 90% between the two numbers?

L: What is it?

S7: 1.645.

L: How about 99%?

S7: 99%, I am not familiar with it.

S8: Which one do you think is the answer of 90%?

S7 pointed at .0505 and .0495: These two, because we should look at their middle, and it's 1.645.

S8 circled .0505 and .0495 and wrote 90% between .0495 and .0485: 90%.

L: How about 99%?

S8 pointed at S7's table and circled .0051 and .0049.

S7: Perhaps we are looking at different places, 2.575…, it should be between 07 and 08. the answer is… 1.57.

S8: Is it?

S7: I think so….1.7…, is it? Is it 50?

L: What's your answer?

S7: 0.050.

L: What's the z_α score?

S7: 2.575.

L: Why is it 2.575?

S7 pointed at the three memorised z scored written to one side: Isn't it?

L: Don't look at them, why do you think it is 2.575?

S7: .0050 is in the middle of the two numbers, so I think it is 2.575.

L: Ok, where is S8?

S6: She went out to answer her mobile phone.

L: Ok. Let's finish the last question.

Q16. Do you know what confidence interval means? What is the difference between 0.95 and 0.90 confidence intervals? Explain in your language.

S6: It's a kind of estimation… to estimate between 2 numbers, and if out of the range, the estimation is wrong.

L: Ok, thanks.

S7: I don't know. What is confidence interval?

L: It's the point estimated value $\pm$ a distance.

S7: I see, it is the value between the calculated range between here and there, it cannot be less or more than the range.

L: Thanks, S8?

S8: It's the value that we want, and it cannot be more or less. It's between the range.

L: What's the range for?

S8: The range is where it may happen.

L: Thanks, S9?

S9: In the graph, the area is 1. I mean, within the value of 1 is the interval.

L: Within the value is the interval… what do you mean the value of 1?

S9: 1 is the graph. The graph equals 1. So 1 may be one of the values in the table.

L: Ok. S10, what is confidence interval?

S10: Confidence interval is the upper limit and lower limit of estimating the number and it will be within this range. It won't be outside.

L: Ok, the last question, what's the difference between 95% confidence interval and

99% confidence interval?
S10: The “value of error” may be smaller.
L: Which one is smaller?
S10: 99%.
L: What if 95% and 90%?
S10: Then the value of error of 95 is smaller.
L: What do you mean its value of error?
S10: Its range of floating.
L: Any other ideas?
S7: I can only see that 95% corresponds to a confirmed number 0.025, but 90% and 99% correspond to ranges between two numbers.
L: You mean when using the data.
S7: Yes, we are shallower.
L: Don’t say that. S8?
S8: No.
L: That’s fine, thank you all. This interview is finished.

You can check the answers now. The answers are 0.488, then 0.8888, 0.999, 0.0202, 0.488 - 0.008 = 0.48, 0.9898 – 0.0808 = 0.909 and the answer of Q (7) is 0, who is correct in Q (7)?
S6: I am.
S7: Me too, I just guessed.
L: Then 3.25, and the last one is -2.5.
S6: Oh, am I a fool? Why did I write this?
[S6 wrote 31.821 - 2.5]
S10: The last question only needs to use “1 –”.
L: What is that?
S6: This is a totally reversed mistake.
L: It only needs “negative”.
S6: I wanted to use “1 -”
L: It doesn’t use “1 -”
S10: Only make it negative?
L: Ok, let me show you by graph. The shadow in the table of t-distribution is in the right-hand side, right?
S6: Yes.
S7: Are the total areas 1?
L: The shadow is in the right-hand side, and the graph looks like this.
S6: Oh! It’s negative.
L: The left-hand side area equals 0.01, and the symmetric positive area is also 0.01.

What we want is the negative one, so it is “minus” two point five.

S6: I see.

S7: Wow.

L: That’s it. Thank you so much today. See you tomorrow.

[This interview took about 90 minutes, and the interviewees had to leave for classes.]

6.8. In vivo codes in the first student interview.

(IN VIVO) CODES	OPEN CODES	CODES
The wrong one should be me	Confidence	1
S2 asks S1	Classmate asking	2
You have to see…, it's the main point	Key idea	3
Not concentrate in class	Class attitude	1
I don't know how to deal with...	Learning result	4
I can only answer the positive…	Selective learning	4
Because I cannot subtract	Reason of difficulty	4
I cannot do questions of Z between 2 values	Difficulty	4
Middle area	Key idea	3
When z = 0, there is no > or <, so it is difficult to guess which side the frequency is.	P (Z = a) = 0, for any a; Not built	4, 5
I think bigger than subtract smaller than becomes equal… the middle 'equal to 3' is 0. each of to sides are 0.5	Area	6
Graph is useless	Useless graph	7, 9
0 is in the middle column, if 3 is in its left-hand or right-hand side, the answer would be 0.03 or 0.05	Number line; Normal curve	4, 10
Negative probability	Negative probability	11
If cannot directly find in the table	Not in the table	12
Ask to remember Z scores corresponding to confidence interval	Memorise	13
Symmetry to think of the other side	Symmetry	3
Not teach difficult content	Degree of difficulty	14
Not concentrate in class (S2 SAID THIS FOR SEVERAL TIMES, SHOULD I REPEAT?)	Class attitude	1
Won't use the table to do business	Practical use of table	9
Drawn pictures make the questions more easily understood	Use of graph	9
I am confused of bigger than and smaller than	Confuse	4
Graph makes it easier to understand	Use of graph	9
Mark the number on the graph, and it becomes clear	Function of graph	9
Not concentrate in class	Class attitude	1
Easier to think of answer, less likely to be wrong.	Function of graph	9
I didn't draw…I use the graph in the front content	Not draw but use graph	7
Memorise answer…memorise formula	memorise	13
Power point and blackboard	Visual tool in class	14
Not concentrate (S1)	Class attitude	1
I don't draw any graph. I use the examples to find how to solve the questions.	Solving strategy Follow example	7
Pure numbers are chaotic… with graphs, numbers	Numbers;	15, 9

are not so chaotic, I can better know which number to put in which place.	Locating	
Never learn statistics before	Back ground	1
I prefer large area (more probabilities)	Table A	16
Two types are the same, but smaller area may confuse me.	Confuse	16, 4
Table A is easier for negative z	Why Table A	16
I am not good in negative z	Difficulty	4
S2 did not notice that the principles may be influenced by > or <	Difficulty of principle	4
Transfer < to >	Transfer	17
If the questions ask positive and I look at the table of negative	Wrong table	4, 16
May forget to turn the page	Page setting	4, 19
What is normalisation?	Don’t know the name	18
It is only a formula for me	Formula for unknown	20
To get a fixed unit, can be used in business or cost	Practical use of normalisation	21
Is it (z score) in this (t) table?	Wrong table	4, 16
0.05 divided by 2 is 0.0025	Mental calculation	4, 10
0.96 corresponds to 95%, and 99% corresponds to 2.757	Wrong memories	4, 13
It’s easy to remember the z scores, so I didn’t think of how they are found.	Never consider	13, 22
Teacher mentions often, but I never remember	Attitude	1
Compare products	Use of confidence interval	23
95% and 90% confidence intervals… 95% mat be bought by more people	Interpretation of confidence interval	5
It is more precise… because it is closer to 100%	More precise	4
Wider range? More choices?	Interpretation of confidence interval	5
Incorrect calculation, other mistakes	mistakes	4
I take advantage of graph	Graph user	7
The answer is the shadow area	Shadow area	6
Why table B? Teacher emphasise table B.	Choice of table	16
I don’t understand the formulas	Difficulty of principles	4, 17
I use table A	Table A	16
Ignore < or >, Ignore +or -, directly use the table	Mistakes	4, 12, 25
Teacher did not emphasise P (Z = a) because it is rare	Not emphasise	14
Graph did not match	Unmatched graph	4, 7, 12
Look at 0.1 when looking for 0,01	Mis-interpreting	4, 12

I only use the table	Only use table	8
Visualisation is drawing graphs.	Definition of visualisation	24
Visualisation include what you see though your eyes, may not be a graph, but a formula or a theorem	Definition of visualisation	24
For some formulas, no matter how hard you try to remember, you may still forget it,…I wish to understand the formula by using graphs, because our brain can remember graphs better, I want to remember the formulas by remember the graphs… I don't rely on memorising formulas.	Why graphs, not memory	13, 7, 9
Visualisation is necessary here but I don't understand.	Need but unable	9, 8
Graphs make the ideas clear	Use of graph	9
It's easier to remember the graphs and not forget.	Why graphs	9
The other teacher transfers the question to a easily-dealt-with form	Easily deal with	16, 17
I understand but don't use table A	Not table A	16
I am not clear with the opinion of > and <. With graph, I have a better understanding, but I still need some time to think of it.	Confused of > and < Use of graph	4, 25
I can also transfer P (-2.41 < Z < -0.03) to P (Z < -0.03) – P (Z < -2.41), I use this method before, but I often make mistakes… because I have to consider which one is bigger…so I use teacher's principle to transfer it to P (Z > -2.41) – P (Z > -0.03)	Alternative method; Often mistake; Smaller and bigger numbers	4, 16, 26
You have to think whether a or b is bigger in P (a < Z < b)…?	a < Z < b, and a > or < b?	4, 10
I can directly use table A when facing negative or smaller than questions. I have to transfer bigger than negative to smaller than or 1 – positive then use table B	Directly use Principles	16, 17
I feel difficult when it is not in the table	Not in table	4, 12
I was not sure if I was right, so I did not write it down	Confidence	1
Is normalisation regrouping data? … It will become positive or negative	Use of normalisation	21
I memorise the z scores	memorise	13
I think I remember the z scores	memorise	13
0.95 confidence interval is wider than 0.90 confidence interval	wider	23
It should be minus 3.1	Mis-read	4, 12
I am not careful enough	Not careful	1, 12
Do you mean how I calculate them? (to transfer)	Transferring is to calculation	17

There is no negative z score in table B	No negative part	27
The teacher did not explain where the formulas come.	Taught but unknown	1, 4
Only taught how to use the table, but how to get any value not in the table was not taught in the 1st semester?	Only table, no principles	14
Do the different representations of Z and t distribution cause your difficulty?	Different representation	1, 4, 12, 28
There are graphs drawn above the tables so I directly use them	Directly use the drawn graph	7
How teachers solve a question is then used by us to solve the next question without thinking of using or drawing graphs. We imitate methods used in the previous question without really understanding.	Imitate; Follow examples; Graphs not needed	26
In fact, I know how to draw graphs but …	Know how but not do	7/8
Questions can be solved by calculation.	Calculation is sufficient.	8
Graphs becomes important in the chapter of probability, there are many events… drawing graphs help explain the ideas. (to make condition clear, but for later part he only needs formula)	Graphs important for events	9
I may be unable to understand the graphs drawn by myself.	Unable to understand self-draw graph	4, 7, 12
I don’t know how to draw a graph for a new question	Draw graph by self	4, 7, 12
Graphs help solving, but itself is difficult to understand	Graph is also a problem	4, 7, 12
I prefer type B… because there is no negative symbol and it looks easier to calculate…I feel it is easier to calculate without negative symbol.	Choice of table; No negative, easier to calculate (negative is difficult)	16, 27, 12
People have difficulty in interpreting or making the graph, they feel it is troublesome.	Difficulty in graph	4, 8, 12
Many people may not notice the existence of the 2nd page of table A.	Page arrangement	4, 12, 19
Without the graph, we can still use other methods to find the answer… I feel that graph is not necessary especially when people know how to calculate with the formulas	Graph not needed	8
I don’t know what the name normalisation means	Don’t know name	18
Normalisation offers a standard.	Practical use of normalisation	21
We rarely notice how the graphs are drawn, and only directly use the formula to calculate.	Graph not focus; Only formula	8, 28
95% confidence interval corresponds to 1.96… 90%	Corresponding;	4, 12,

is near by 95%, so I think 95% will be nearby…1.945?... 0.025 correspond to 1.96, so I think 0.5 will be close to it… When looking for 0.05, S5 pointed at .05 on top z score, not inside probability	Multi-misinterpreting; Confused of outside; z scores and inside; probabilities	13, 29
I have to search one by one	Unaware of tendency	29
They are all 0.05 (.0548, .0537, .0526, .0516, and .0505)	All 0.05	4, 12, 10
S5 tried again to find 0.05 in table from up to bottom, rather than from left to right	Search approach	16
L: Which is closest to .05? S5 said .0505. L: Anything else? S5: I don't know S5 wrote 0.0505 = 1.64	Constrained by condition .05; Never consider 0.4995; Meaning of =	10
When you calculate 1 – 0.95, its confidence interval is located at 0.05	[S5 perhaps meant 0.05 is the area of two sides. He know how to proceed.]	3
90% is 1.645, 95% is 1.96, and 1.96 is longer than 1.645, so 95% is longer than 95%	Length of confidence interval	23
I drew graphs to realise how to get the answer.	Graph user	7, 9
I decide which table to use according to whether the question asks > or <, but in most questions, I use table B. if the question asks <, I will use 1 – to become >… I would use table A to check my answer of questions involving with negative symbols.	Choice of table; Principles; Table B	16, 25, 27
I make decision according to > or <… I use type B if there is a negative value, I will write 1 – positive value.	Principle; Table B	16, 25, 27
I also make decision according to > or <. I use type B and the negative part of type A.	Principle; Table B and negative A	16, 25, 27
I feel table of z is easier to search than table of t, because I have more impression of it	Familiarity	30
I don't know where 0.99 is in the table of t distribution… graph does not help…I think I can get the answer from 1 – the answer of 0.01 (0.99 = 1 – 0.01, so the answer of 0.99 = 1 – answer of 0.01)	Not in the table; Unmatched graph; Mis-corresponding	4, 7, 12, 31
I feel t distribution is easier than z, but I have to consider v… Teacher gives 0.xx followed by a number, and it is easier to find. (Probabilities in the table are rather small, and questions and examples do not ask probabilities close to 1, so S7 did not know	Question of z may require graph, so difficult	14, 22

such question, and thought values always can be found in the table)		
Remember the principles and judge	Principle	17
The other one of 0.99 is 0.01, so I directly look at 0.01 and add a negative symbol.	Symmetry	3
I did not know that I should directly at the v given in the question. I remember that there is a '- 1' to get v	Representation V = something -1	4, 12
If the number is exactly on the point, I can find it. When there is a negative symbol, I will use 1 – the number corresponding to the positive one in the table. But in this question, 1 – 0.99 = 0.01, so 1 – the number corresponding to 0.01 would be the answer of 0.99. (Follow the same principle used in table of Z, but t is in a reverse format.)	Exactly on point; Mis-corresponding; Follow methods.	27,17, 28, 31
I didn't write because I only guess it.	Confidence	1
S8: visualisation is… I realise it when I see it. S6: you realise it or hate it?	Definition of visualization; Realise or hate	24, 1
When I see graphs, I know they are difficult right away, so I hate graphs. But I still draw graphs when solving questions.	Hate but still draw	1, 7
I like some graph (charts for data representation), and hate some graphs of stand normal distribution that requires judgment. Graphs of set are necessary, I don't like them but I still draw them to understand. However, I don't like graphs of normal distribution, and I don't want to learn it. You may use graph to solve question, but you can also solve the questions without drawing graphs, so I feel I don't have not learn it.	Like graphs; Don't like but draw; Don't like and don't want to learn; Graphs not needed	1, 7/8, 9
I feel that some graphs let me have better understanding of what it is talking about than words do, such as intersection and union of sets. Graphs in latter chapter are more difficult, and it may be influenced by my interpretation of teacher's speech in class.	Better understanding than words; Interpretation of teacher's speech	9, 1, 14
I don't know graphs because I did not memorise how to draw and interpret graphs. I only remember how to solve by using formulas. I memorise them when I write them.	No graph; Only Formulas; Solving orientation	8, 17, 32
If I see a graph, I will just ignore it. I will use my own method (not explain)	Just ignore graphs; Unknown alternative method	1, 8
S6 hated but used graph. S8 does not hate but not learn graph.	Hate but use; Not hate but not learn	1, 7, 8
Using graph…depends on which chapter it is, and	Depends on chapter;	1, 7

depends on my mood	Depends on mood	
I draw when necessary, but not for normal calculation.	Graph not for calculation	1, 7
I didn't memorise the answers, because I may lose some process in my memory.	Memorising may lose process.	13
There are 2 kinds of graphs: charts in earlier part and graphs for distribution which requires calculation and making decision. I like charts. Latter graphs are more difficult and I don't understand, so I just let it be.	Let it be	1, 7/8
I don't draw graph because it is troublesome. (Some are easy) I just don't want to	No graph	8
Sine and cosine cannot be understood without graph. (the ideas comes form shape of triangle)	Must graph	9
I can easily see which is/are more or fewer in a peek if it is in a bar chart. Without graph, you have to count each one of them.	Aware of tendency in a peek	9
The other teacher explain they are table of bigger than and table of smaller than	Interpretation of table A and table B	16
I prefer table B, but if the answer of asking smaller than can be directly found in table A, I will use it without troublesome transferring…the important thing is convenient to use.	Troublesome transferring; Important thing: convenient	17
I don't use table A often and the numbers are all negative, so it is more difficult to judge. I am not familiar with this table.	All negative, more difficult; Unaware of positive part of table A	4, 12, 27, 30
It is easier to find if it is on the point.	Exactly on the point	33
I don't prefer any table, but I think table A is easier to find, because it includes both negative and positive, so the questions should be more likely included and the answer can be directly found in the table.	Table A easier to use; Includes both negative and positive (more) values	16, 27
I prefer table B because I don't bother to turn the page	Page arrangement	16, 19
I use whichever table I got… no difference.	whichever	16
I feel the only question is there are too many numbers in the both tables	Too many numbers	4, 12
When I use table B for question asking smaller than or negative, I have to decide to use '1 –' or transfer negative to positive.	Judgment	16, 17, 27
It is difficult to find if it is not on the point, I have problem when it involves transferring.	Not exactly on the point; Difficulty of principle	4, 33, 17
I may find a neighbour.	Mis-targeting	12

The words are too small, and I may find a wrong one or find a correct one but write incorrectly.	Small words	12
Purpose of normalisation is to use the table.	Purpose of normalisation	21
I don't know the purpose of normalization. I learn it because it is taught.	Learn because taught	1
I don't know the meaning, and only know that the answer can be gotten by calculation.	Solving orientation	33
I memorise the formula of normalisation… I don't know how it is generated.	Memorise without understanding	13, 4
The formula is taught like this, so I write it like this	Write what is taught	1
Where can I find it? (when asked to find the z score corresponding to 95% confidence, some students such as S8 did not know to do (1 – 0.95)/2,)	Never consider	4, 12
Why is there such a question, I can memorise them, 90% is 1.575	Never consider; Wrong memories: 90% is 1.645 and 99% is 2.575. there is no 1.575	4, 13
2.44, I also remember them	Wrong memory	13
S7 murmured: 1.96… L pointed at .0197: are you looking for 1.96? S7: to look for two close numbers. S7 pointed at .1977: here! Oh, no… it is too difficult. (S7 remember 1.96 corresponds to 95%, but she could not find 1.96 in the probabilities, because 1.96 is outside z score, and she was looking at inside probabilities)	Close numbers; Inner/outer of table	4, 13
I cannot find anything about that	No idea	4, 22
Later S6 found the answer: I really understand it now, it is so impressive	Really understanding of impressive difficulty	22, 34
S9 wrote 0.01/2 = 0.005, and look for it in the table. Initially he pointed at .0005, but moved away and stop at .0051 and .0049 (corresponding to 2.57 and 2.58): the answer is 2.578… He wrote .005 between .0051 and .0049 and said 0.05 is here, the left side is 2.5 and top 78. Soon he corrected his answer to 2.575.	Mis-interpreting; Mis-interpreting	12
For 95% confidence interval, S7 found z score 1.9 and .06 corresponding to .025, but did not know what it is. S10 explain it involves the possibility of positive and possibility of negative, so 0.05 should be divided by 2.	Memorising z scores without/with involving $(1 - \alpha)/2$	12, 13
S7: how about 99%?	Mental calculation	10

S6: 0.01 divided by 2. S10: 0.025 S7: 0.025 S8: should be 0.05		
S10 suggest others to look for z score corresponding to 90% confidence interval S8 tried to find 0.05 in the table, she stopped at .0051 and .0049, but soon move away. Later S8 pointed at .0052 and .0051: here L: what is your answer? S8 pointed at .0051: this one.	Aware of misinterpreting; Unaware of between	12, 22
S7: why is there 1.575? (Perhaps she remembered 1.575 corresponding to 90%)… I found 1.96 in this way, but where is 1.575 (she did not consider (1 – 0.9 /2) and only tried to find the 1.575 mistakenly remembered in her memory.) I am not familiar with this Later S7 pointed at .0505 and .0495: is the answer of 90% between the two numbers? L: what is it? S7: 1.645 S8: which one do you think the answer? S7 pointed at .0505 and .0495: we should look at their middle	Mixed and wrong memories: (2.575 corresponds to 99% 1.645 corresponds to 90%); Influenced by image Unfamiliarity Find between	13, 4, 12, 30, 35
L: how about 99%? S7: 2.575… because 0.005 is in the middle of the two numbers 2.57 and 2.58.	between	35
Confidence interval is a kind of estimation, to estimate between two numbers, and if out of the range, the estimation is wrong	Interpretation of confidence interval	23
Confidence interval is the value within the calculated range from here to there, and it cannot be less or more.	Interpretation of confidence interval	23
Confidence interval is the value that we want, and it cannot be more or less. It is within the range… the range is where it may happen.	Interpretation of confidence interval	23
In the graph, the area is 1… within the value of 1 is the interval…the graph equals 1. So 1 may be a value in the table.	Interpretation of confidence interval	23, 5
Confidence interval is the upper limit and lower limit of estimating the number, and will be within the range. It won’t be outside.	Interpretation of confidence interval	23
The difference between 95% and 99% confidence interval is that value of error of 99% is smaller. Value of error is the range of floating.	Value of error	23, 5
S7 is correct in Q7: I just guess.	Guess	1
Why did I write this?	Couldn’t understand	22, 27

Oh! It’s negative.	what he wrote 1 hour ago; Never consider; negative.	

6.9. Codes in the first student interview.

NUMBER	CODES	INCLUDE	BELONG TO	LINK TO
1	Student issues	32	12	
2	Students’ interaction			
3	**Key idea**	6, 17, 25, 27, 33, 35		
4	**Difficulty**	17, 25, 27		12
5	Self theory		12	30
6	Area		3	
7	**Graph user**		36	
8	**Graph non-user**		36	
9	**Function of graph**		20	
10	Fundamental knowledge		12	
11	Disciplinary mistake		5, 12	
12	**Cause of difficulty**	5, 10, 13, 15, 19, 22, 30, 31		4
13	**Memorise**		12, 36	
14	Class issues		12	
15	Numbers		12	
16	Choice of table		36	
17	**Transfer/Principle**		3, 4	
18	Name		30	
19	Page arrangement		12	
20	Interpretation	9, 21, 23,24		
21	Function of normalization		20	
22	**Never consider**	27	12	
23	Function/interpretation of confidence interval		20	
24	Definition of visualisation		20	
25	> **and** <		3, 4, 12	
26	**Follow**		36	
27	**Negative**		3, 4, 12, 22	
28	Format		12	
29	**Inner/outer of table**		12	
30	(Un)familiarity	18	12	5
31	**Mis-corresponding**		12	
32	**Solving orientation**		1, 36	
33	(Not) on the point		3	
34	Impressive			
35	Between		3	
36	**Students’ strategy**	7, 8, 16, 17, 26, 32		

6.10. Categories in the first student interview.

CATEGORIES	INCLUDING CODES	SUCH AS
Key Ideas (Must-know & Factor)	Area	
	Transfer/Principle	
	> and <	
	Negative	
	(Not) on the point or table	
	Between	
Difficulty & Cause of Difficulty	Transfer / Principle	
	> and <	
	Negative	
	Not on the point	
	Format	
	Inner/outer of table	
	Student issues	Attitude; Experience
	Fundamental knowledge	Mean; Variance & S.D; Empirical rules
	Numbers / Number line	
	Mis-corresponding (homography)	
	Mistakes of using the table	Mis-reading; Mis-transcribing; Mis-transferring; Mis-targeting; Mis-interpreting; Mis-calculation
	Memory	
	Unfamiliarity	
	Self-theory	
	Never consider	$\neq$; Negative
	Class issues	Teacher's mistake
	Material (textbook)	Page arrangement; Incorrectness; Content

Students' Strategies	Graph users	(un)matched graphs
	Graph non-users	
	Choice of tables	
	Transfer/Principles	
	Follow	
	Memorise	
	Solving orientation	
Students' Interpretation	Definition of visualisation	
	Function of graph	
	Function of normalisation	
	Function/definition of confidence interval	
Others (Rarely Mentioned)	Students' interaction	
	Impressive of the process of realising	
	t distribution	

6.11. Questions of the second student interview.

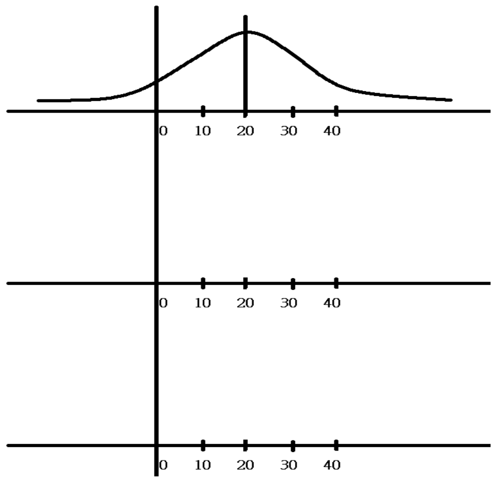
Part I

Q1. P (Z > a) = 0.67; P (Z > 0.67) = b; P (Z < c) = 0.67; P (Z < 0.67) = d; and P (Z = 0.67) = e. Find a, b, c, d and e.

Part II

Q2. Get two independent random samples (X1, X2,,…, X20) & (Y1, Y2,…, Y20) from two normal populations X ~ N (20, 20x20) and Y ~ N (10, 10x10) respectively. The sampling distribution of (X bar – Y bar) follows N (a, b). What are a and b?

Q3. Following Q2, X ~ N (20, 20x20) is shown on top, please draw Y ~ N (10, 10x10) and (X bar – Y bar) ~ N (a, b) in the second and third places.



*The graph with scales was only given to S6-S10

Q4. Following Q2, what is the 95% confidence interval of (X bar – Y bar)?

Q5. Following Q2, are the two populations significantly different?

Part III

Q6. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c.

Please test whether the content is 500c.c. and write down your steps. $\alpha = 0.05$

Q7. It is said by a drink seller that the content of their drink is over 500c.c. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10c.c. Please test whether the drink seller is telling the truth, and write down your steps. $\alpha = 0.05$

Q8. Which method did you use in Q6 and Q7? Why did you use this method?

Q9. Did/Can you draw graph for Q6 and Q7?

6.12. The second student interview with S1 and S2 on 1st Jun 2009.

[In the second student interview, I asked S1&S2, S3&S4, S5 and S6-S10 to solve/answer nine questions. The interviews took place two or three weeks before the final exam, and many of them had not studied yet, especially the chapter of hypothesis testing. I tested Q6 and Q7 again in the third interview with S1-S5 whereas S6-S10 could not participate after the final exam finished.]

L: Hello, S1 and S2, thanks for coming today. This time I prepared nine questions for you, I would like you to solve them. The first question is testing how you understand the table of normal distribution. Please start solving the 4 sub-questions.

Q1. P (Z > a) = 0.67; P (Z > 0.67) = b; P (Z < c) = 0.67; and P (Z < 0.67) = d. What are a, b, c, and d?

[In the second sub-question, S2 moved his finger in the positive page of type A and stopped near P (Z < 0.45) = .6736 but soon moved away, and stopped at P (Z < 0.44) = .6700 and moved away again. He spent more than 3 minutes and still looked confused.]

L: Which question are you solving?

S2: Second one. I don't understand.

L: There are 3 pages involving standard normal distribution. You can keep trying or solve the third one first.

S2: Ok.

[S2 turned to the negative part of Table A, found nothing, turned back to the positive part, and soon turned to the negative part again. Twelve minutes after the questions were given. S2 gave up answering, whereas S1 finished.]

Q1	**P (Z > a) = 0.67**	**P (Z > 0.67) = b**	**P (Z < c) = 0.67**	**P (Z < 0.67) = d**
Correct answer	**a = -0.44**	**b = 0.2514**	**c = 0.44**	**d = 0.7486**
S1 (type A and B)	a = -0.44	b = 0.2514	c = 0.44	d = 0.7486
S2 (type A)	0.44			
S2 supplemented (type A)		~~0.2206~~ 0.2514	1 - 0.67 = 0.33 c = 0.6293	

L: Ok, it's not a formal test. I just want to know how you think when you use this table. May I ask how you find 0.44 in the first sub-question?

S2 pointed at .6700 with .4 in the left and .04 on the top: here.

L: Where does it come from?

S2: Use the table, I found 0.67, and then correspond to .4 and .04.

L: I see. The third sub-question is also 0.67, what is c?

S2: Is it 0.23?

L: Why 0.23?
S2: Because it is <, so 1 – 0.67
L: 1 – 0.67? It should be 0.33
S2 laughed: Oh, yes.
L: That's fine, please write it down, and try if you can to find the answer?
S2 pointed at 0.6293 with .3 in the left and .03 on top: this one.
L: Ok, please write it down.
S2: Ok, 0.6293.
L: How about the fourth sub-question?
S2: Smaller than 0.67….. I don't know.
L: And second one?
S2: No.
L: That's fine. Do you use type B?
S2: No.
L: Thanks. S1?
S1: Yes.
L: Which table do you use?
S1: Both.
L: Ok, how do you find a?
S1: Because 0.67 can be found in the table [i.e. Table A], so I directly find it. But the question asks "bigger than" and this table is "smaller than", so it is negative.
L: Ok, where did you find 0.44?
S1: Here .4 and this .04 correspond.
L: Ok, how about b?
S1: B, asks you to find 0.67. In the table of type B find .6 and .07 and directly get the answer.
L: In the table of type B, 0.2514?
S1: Yes.
L: Ok, the third sub-question?
S1: The answer of "smaller than" c is 0.67, so directly look at the Table A, and this time it is "smaller than", so the answer is 0.44.
L: Ok, the fourth sub-question?
S1: "smaller than 0.67", directly look at the table of type A, .6 and .07. Just like this.
L: 0.7486, correct. Do you understand now, S2?
S2: Hmm…
L: In your first and third sub-questions, you used a reverse approach of using the table. You can discuss with S1 after this interview. Next part is the content taught after the mid-term. Perhaps you did not study yet, but you can try if you can answer the second question.

Q2. Get two set of independent random samples ($x_1, x_2, ... x_{20}$) & ($y_1, y_2, ... y_{20}$) from two normal populations X ~ N (20, 20^2) and Y ~ N (10, 10^2) respectively. The sampling distribution of ($\overline{X} - \overline{Y}$) follows N (a, b). What are a and b?

L: The content is in Section 8-6, and you can use your textbook and notes.

S1: Should I calculate them?

L: Yes, please.

5 minutes later, S1: This does not involve confidence interval, does it?

L: No. it doesn't. This question asks the mean and variance of $\overline{X} - \overline{Y}$.

10 minutes after the question is given, S1: I give up, my brain is rusty. At this moment, S2 wrote Z =…= $\dfrac{(\overline{X} - \overline{Y}) - 10}{5}$ ~ N (0, 1).

L: Ok, the content is in the page 279, would you like to try again?

2 minutes later, S1: Aren't they 10 and 25?

L: Yes, they are easy, right?

S1: Oh, my god.

L: Let me tell you something. There is a problem on page 279 in your textbook: it's mistakenly printed that $\overline{X} - \overline{Y}$ ~ N ($\mu_1 - \mu_2, \sqrt{\dfrac{\sigma_1^{\ 2}}{n_1} + \dfrac{\sigma_2^{\ 2}}{n_2}}$), but as you can see on page 280, it should be $\overline{X} - \overline{Y}$ ~ N ($\mu_1 - \mu_2, \dfrac{\sigma_1^{\ 2}}{n_1} + \dfrac{\sigma_2^{\ 2}}{n_2}$). The place after mean is not standard deviation, but variance or square of standard deviation. So a is 10 and b is 25 or 5^2, not 5.

	Q2
Correct answer	$\mu_1 = 20, \sigma_1 = 20, n_1 = 20; \mu_2 = 10, \sigma_2 = 10, n_2 = 20$ $\overline{X} - \overline{Y}$ ~ N ($\mu_1 - \mu_2, \dfrac{\sigma_1^{\ 2}}{n_1} + \dfrac{\sigma_2^{\ 2}}{n_2}$) = N (20 – 10, $\dfrac{20^2}{20} + \dfrac{10^2}{20}$) = N (10, 25) a = 10, b = 25
S1	$\overline{X} - \overline{Y}$ ~ N ($\mu_1 - \mu_2, \dfrac{\sigma_1^{\ 2}}{n_1} + \dfrac{\sigma_2^{\ 2}}{n_2}$) =?

	$n_1 = 20$, $\overline{X}$ =? s_1 = 400 n_2 = ~~10~~ 20, $\overline{Y}$ =? s_2 = 100 $(\overline{X} - \overline{Y}) \times \sqrt{\frac{400}{20} + \frac{100}{20}} = (\overline{X} - \overline{Y}) \times 5$
S1 told p.279	$\mu_1 - \mu_2$ = 20 – 10 = 10 $\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}\right)$ $\overline{X} - \overline{Y} \sim (10, \sqrt{25})$
S2	N (20, 400), N (10, 100) $Z = \frac{(\overline{X} - \overline{Y}) - (20 - 10)}{\sqrt{\frac{400}{20} + \frac{100}{20}}} \sim N(0,1)$ $= \frac{(\overline{X} - \overline{Y}) - 10}{\sqrt{20 + 5}} \sim N(0,1)$ $= \frac{(\overline{X} - \overline{Y}) - 10}{5} \sim N(0,1)$
S2 told p. 279	$\overline{X} - \overline{Y} \sim N\left(20 - 10, \sqrt{\frac{400}{20} + \frac{100}{20}}\right)$ $\overline{X} - \overline{Y} \sim N(10, \sqrt{25})$ $\overline{X} - \overline{Y} \sim N(10, 5)$

Q3. Following Q2, X ~ N (20, 20^2) is shown on top, please draw and Y ~ N (10, 10^2) and $\overline{X} - \overline{Y}$ ~ N (10, 5^2) in the second and third places.

S1 was thinking but S2 was uninterested in this question.

L: S2, don’t you want to try?

S2: I don’t know how to draw.

L: That’s fine. S1, what do you write?

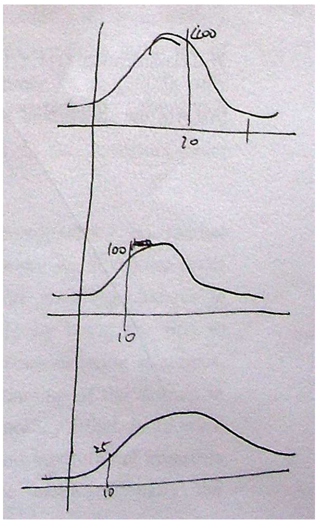
S1: 100.

L: What does 100 mean?

S1: Variance.

L: Why do you write 100 over there?

S1 drew 2 similar curves like mine, added 2 vertical lines and numbers in the graph; S2 gave up.



L: Ok, why is the third curve wider than others?

S1: Because its mean is 10, but its variance is smaller.

L: Wider because of smaller variance?

S1: Right, and its beginning part is flatter. It's not as sharp as the second curve.

L: Its variance, 25, is smaller, so the curve is flatter. Is that what you mean?

S1: Yes.

L: Ok, thanks.

L: I saw that S1 just used the formula to get confidence interval in Q2.

S1: Oh yes.

L: You can do it now.

Q4. Following Q2, what is the 95% confidence interval of $\overline{X}-\overline{Y}$?

S1: Should I calculate to get the answer?

L: Yes, please.

S1: Then I need a calculator.

L: Go ahead. S2, can you solve this?

S2 wrote on paper: hmm… I can only do this.

L: Ok. S2, where is the formula you use?

S2 pointed at $\left[(\overline{X}-\overline{Y}) - Z_{\frac{\alpha}{2}}\times\sqrt{\frac{\sigma_1^{2}}{n_1}+\frac{\sigma_2^{2}}{n_2}},\ (\overline{X}-\overline{Y}) + Z_{\frac{\alpha}{2}}\times\sqrt{\frac{\sigma_1^{2}}{n_1}+\frac{\sigma_2^{2}}{n_2}}\right]$ in page 280: here.

L: Where is the part of square root?

S2: Am I wrong to directly multiply the answer in second question?

L: $Z_{\frac{\alpha}{2}}$ becomes 1.96?

S2: Yes.

L: Why not multiplying?

S2: Should I? Oh, I am sorry.

L: You don't have to.

S1 replaced ($\overline{X} - \overline{Y}$) = 5 by ($\overline{X} - \overline{Y}$) = 10 and rewrote Q4.

L: Ok. Please answer question 5 according to your present answer of question 4.

Q5. Following Q2, are the two populations significantly different?

S1: Only answer yes or no?

L: And your reason, please. I know that teacher A emphasised this.

S1 murmured: One is positive, and the other is negative… difference is small…

S2: I am finished.

S1: Me, too.

S1: It's written in the textbook [P. 287].

L: Do you find it in the book or do you remember what teacher A said?

S1: I remember teacher A mentioned including or not including 0.

L: Ok, do you have answer now, S2?

S2: Yes.

L: Do they have significant difference?

S2: Yes.

L: Why?

S2: Not including 0. Am I right?

L: Thanks.

	Q4	**Q5**
Correct answer	$\frac{1-0.95}{2} = \frac{0.05}{2} = 0.025$ $Z_{0.025} = 1.96$ **(** $\overline{X} - \overline{Y}$ **) ±** $Z_{0.025} \times \sqrt{\frac{400}{20} + \frac{100}{20}}$ **= (20 - 10) ± 1.96 × 5 =10 ± 9.8 = (0.2, 19.8)**	**They are significantly different because 0 is not included in (0.2, 19.8)**
S1	1 – α = 0.95, α = 0.05, $Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$ $\cancel{(\overline{X} - \overline{Y}) \times 1.96 \times \sqrt{\frac{400}{20} + \frac{100}{20}}}$ ~~= 5 ± 1.96 × 5~~ ~~= (5 + 9.8, 5 – 9.8)~~ ~~= (14.8, -4.8)~~	Significantly different. Not include 0.

	$(\overline{X}-\overline{Y})\times 1.96\times\sqrt{\frac{400}{20}+\frac{100}{20}}$ = (10 + 9.8, 10 – 9.8) = (19.8, 0.2)	
S2	(5 – 1.96, 5 + 1.96) (3.04, 6.96)	Yes, μ1-μ2 does not include 0, so there is a significant difference

Q6. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the content is 500c.c. and write down your steps. α = 0.05

L: Please don’t just guess.

S2: The claim is real because the standard deviation is 10c.c. and the difference of 498.2 and 500 is only 0.2c.c. Thus 100 bottles should be within it.

L: Can you write it down?

S2: Ok.

S2 wrote.

L: Can you write down hypotheses?

S2: …

S1 only wrote hypotheses: what else should I write?

L: The process to test whether the content is 500c.c. Many questions in your textbook only ask the principles of hypothesis testing, but I want you to finish the process.

S1 did not study yet, and spent a few minutes to browse the textbook.

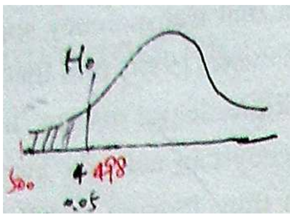
S1: 0.05…

L: α equals 0.05

S1: Confidence interval?

L: No, it’s not confidence interval. It’s called significant level.

S1 drew a normal curve and a left tail, but he was not sure and murmured: this side or that side?



S1 wrote 500 × 0.05 = 25 > 10; so reject.

L: What does 500 × 0.05 mean?

S1: 500c.c. and under the level 0.05, so multiply them, and get 25. It’s 15 bigger than standard deviation.

L: Ok, this is what you think. The next question is very similar to this one.

Q7. It is said by a drink seller that the content of their drink is over 500c.c.. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the drink seller is telling the truth, and write down your steps. α = 0.05

S1: Is the process taught in the textbook?

L: Yes, it's also taught by teacher A, but most examples only ask the testing principles.

S1: Ok, I am done.

L: H1 is confirmed. Is the claim true?

S1: Yes.

L: Why is the claim true?

S1: If the standard deviations $\pm$ 10 add mean, then it will exceed.

L: Exceed what?

S1: 500.

L: I see. Would you like to say something, S2?

S2: No.

	Q6	Q7
Correct answer	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ = 500** H_1 **: μ ≠ 500 (two-tailed testing)** $Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$ **-** $Z_{\frac{\alpha}{2}}$ **= -1.96** **CR = {Z > 1.96 or Z < -1.96}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 > -1.96** **Accept** H_0**, the content is 500c.c.**	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ** $\geq$ **500** H_1 **: μ < 500 (left-tailed testing)** **-Zα = -Z0.05 = -1.645** **CR = {Z < -1.645}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 < -1.645** **Reject** H_0**, the drink seller is not telling the truth.**

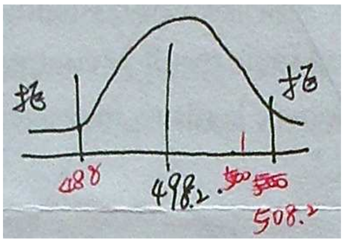
S1	H_0 = 500 H_1 < 490 500 × 0.05 = 25 > 10 So reject	H_0 < 500 H_1 > 500 Standard = 10c.c. Average 498.2c.c. H_1 is confirmed. True.
S2	standard deviation 10; examine and get 498.2c.c. ~~not exceeding standard deviation 10c.c., so they all are.~~ then in the range of standard deviation, So it is true.	μ_0 = 500, n = 100, $\overline{X}$ = 498.2, s = 10, α = 0.05 H_0 : μ > 500, H_1 : μ< 500

Q8. Which approach do you use to solve question 6 and 7? Why do you use this approach?

[S1 and S2 did not answer.]

Q9. Can you draw graphs in question 6 and 7?

[S2 did not draw. S1 drew a graph when he was solving Q6. He drew another graph for Q7 (see below) when Q9 required him to do so.]



L: Can you locate 500 and 498.2 in the graph of Q6?

S1: 500 should be behind 498.

L: And graph for question 7.

[S1 located 488 and 508.2 in the left- and right-hand side of 498.2 which was in the middle.]

L: Why the two numbers?

S1: Because of ±10.

L: Where is 500?

S1 marked 500 between 498.2 and 508.2, but closer to 508.2: About here. Sorry, I didn't learn well.

L: Don't worry. Thank you so much. See you tomorrow.

S1 and S2: Bye.

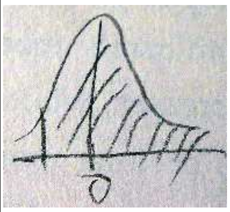
6.13. The second student interview with S3 and S4 on 3rd Jun 2009.

[In the second student interview, I asked S1&S2, S3&S4, S5 and S6-S10 to solve/answer nine questions. The interviews took place two or three weeks before the final exam, and many of them had not studied yet, especially the chapter of hypothesis testing. I tested Q6 and Q7 again in the third interview with S1-S5 whereas S6-S10 could not participate after the final exam finished.]

L: Hello, S3 and S4, thanks for coming today. This time I prepared nine questions for you, I would like you to solve them. The first question is testing how you understand the table of normal distribution. Please start solving the four sub-questions.

Q1. P (Z > a) = 0.67; P (Z > 0.67) = b; P (Z < c) = 0.67; and P (Z < 0.67) = d.
What are a, b, c, and d?

15 minutes later, S3 and S4 put their pen on the desk (that means stop).

Q1	**P (Z > a) = 0.67**	**P (Z > 0.67) = b**	**P (Z < c) = 0.67**	**P (Z < 0.67) = d**
Correct answer	**a = -0.44**	**b = 0.2514**	**c = 0.44**	**d = 0.7486**
S3 (type)	~~1 - P(Z > a) =1 - 0.67~~ P(Z<-a)=1-0.67=0.33 -a = -0.44 a = 0.44  [S3 retried.] P(Z > a) = 0.67 1 - P(Z > a) = 0.67 P(Z > a) = 0.33 a = 0.44	b = 0.2514	c = 0.44	d = 0.7486
S4 (type)	a = 0.2514			

L: S4, you still have difficulty, right?

S4: …

L: That’s alright, you can listen to what I discuss with S3, or you can ask her after this interview if you are interested.

L: S3, you are confused about ‘a’, but can you tell me how you find ‘b’?

S3: B is the area, and 0.67 is α.

L: Right.

S3: So I look for 0.67 in the table, and this should be ‘b’.

L: Right. S4, your answer for ‘a’ is actually the answer for ‘b’. You mixed up the location of Z score and probability so you could not find ‘b’. How about ‘c’?

[S4 did not answer.]

S3: P (Z < c) = 0.67, so c = 0.44.

L: Yes, and ‘d’?

S3: Should be 0.7486.

L: Right. Ok, you only have difficulty in ‘a’, you solve it twice and also draw a graph. I designed the questions by changing “smaller than” and “bigger than”, and changing location of 0.67 and the answers. You cannot directly find ‘a’ in the table. In your answer, P (Z > a) = 0.67, and P (Z < -a) = 1 – 0.67 = 0.33. In fact, P (Z < -a) should be still 0.67, and it equals P (Z > a).

S3: ?

L: You can write it down: P (Z < a) = 1 – P (Z > a). Isn’t it?

S3 laughs: Right.

L: This is the only question which requires “1 -”, and what is the result?

S3: 1 – 0.67 = 0.33.

L: Then you can look for ‘a’ now.

S3: P (Z < a) =0.33, so ‘a’ = 0.44.

L: That’s right. Ok, let’s move on to the second question.

Q2. Get two sets of independent random sample ($x_1, x_2,...x_{20}$) & ($y_1, y_2,...y_{20}$) from two normal populations X ~ N (20, 20^2) and Y ~ N (10, 10^2) respectively. The sampling distribution of ($\overline{X} - \overline{Y}$) follows N (a, b). What are a and b?

L: I designed this question according to an example in the textbook, the content is in Section 8-6, and you can use your textbook and notes.

5 minutes later, they put their pens on the desk.

	Q2
Correct answer	$\mu_1 = 20, \sigma_1 = 20, n_1 = 20; \mu_2 = 10, \sigma_2 = 10, n_2 = 20$ $\overline{X} - \overline{Y} \sim N\,(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$ $= N\,(20 - 10, \frac{20^2}{20} + \frac{10^2}{20})$ $= N\,(10, 25)$ a = 10, b = 25

S3	$\mu_1 = 20, \sigma_1^2 = 20^2, n_1 = 20$ $\mu_2 = 10, \sigma_2^2 = 10^2, n_2 = 20$ a = $\mu_1 - \mu_2$ = 20 – 10 = 10 b = $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ = $\frac{20^2}{20} + \frac{10^2}{20}$ = $\frac{500}{20}$ = 25
S4	N (20, 20^2), N (10, 10^2), 20 $(x_1, x_2, \ldots x_{20})$ $(y_1, y_2, \ldots y_{20})$

L: If possible, can you draw the graphs?

Q3. Following Q2, X ~ N (20, 20^2) is shown on top, please draw and Y ~ N (10, 10^2) and $\overline{X} - \overline{Y}$ ~ N (10, 5^2) in the second and third places.

S3: I am willing to try.

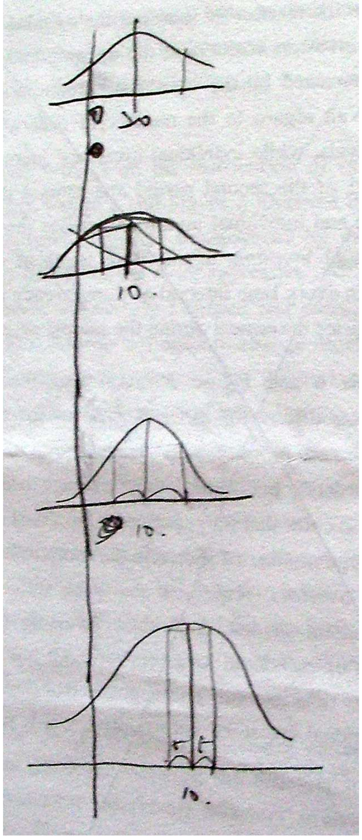
L: Good.

[Three minutes later, they looked confused. Their first interview took a long time, so this time they wished to finish on time, so I could not give them too much time to think about each question.]

L: Ok, as you see the graph of N (20, 20^2), and you just need to draw N (10, 10^2) and N (a, b). Now I can tell you that a = 10 and b = 25 = 5^2. How do you draw them? I will observe your understanding of its shape, size and location.

[Three minutes later.]

L: S3, now you have 4 graphs [S3 deleted her second graph], which one is the second one?

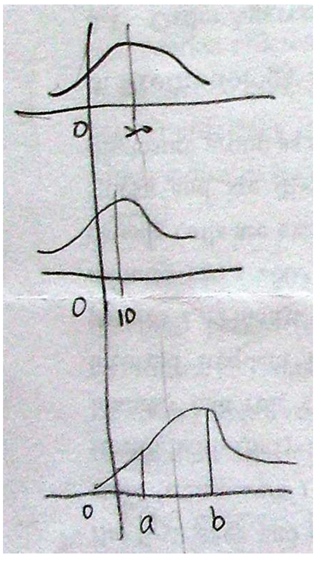


S3 points at third one: This.

L: And the third graph?

S3 points at the fourth graph: This.

L: I see. S4, can you explain ‘a’ and ‘b’ in your graph? I want to know how you think, such as why ‘a’ is here and ‘b’ is there.



S4: It is… ‘b’ is 25, so it is a little bit bigger than 20.

L: So the central line is 20?

S4: Yes.

L: Ok. S3, can you explain your graph.

S3: The second one… this is 10 and that is also 10 [i.e. the distances].
L: Can you mark them?
S3: Ok.
L: Thanks. How about the last graph?
S3: The centre is 10, this distance is 5 and the other distance is 5.
L: I see. I think I should draw them for you. In the first graph, the centre is 20, and standard deviation is 20. According to empirical rules, a fixed ratio would be in the area of (mean – standard deviation, mean + standard deviation), so the centre of the second graph is 10, and it looks smaller. The centre of the third graph is still 10, but its standard deviation is 5, so it is even smaller. I think you didn't have many chances to draw them, so the sizes of your graphs are not reasonable. That's fine, let's see the fourth question.

Q4. Following Q2, what is the 95% confidence interval of $\overline{X} - \overline{Y}$?
L: Let me remind you that $\overline{X} - \overline{Y} \sim N(10, 5^2)$, and you can find it in your textbook.
8 minutes later, L: After Q4, you can also try Q5.

Q5. Following Q2, are the two populations significantly different?
5 minutes later, S3 and S4 did not understand what Q5 is about. Because of the time constraint, I let them stop answering Q5.

	Q4	**Q5**
Correct answer	$\frac{1-0.95}{2} = \frac{0.05}{2} = 0.025$ $Z_{0.025} = 1.96$ $(\overline{X} - \overline{Y}) \pm Z_{0.025} \times \sqrt{\frac{400}{20} + \frac{100}{20}}$ **= (20 - 10) ± 1.96 × 5 =10 ± 9.8 = (0.2, 19.8)**	**They are significantly different because 0 is not included in (0.2, 19.8)**
S3	1 –95% = 5% = 0.05 Confidence interval $= (\overline{X} - \overline{Y}) \pm Z_{\frac{0.05}{2}} \times \sqrt{\frac{20^2}{20} + \frac{10^2}{20}}$ $= (\overline{X} - \overline{Y}) \pm Z_{0.025} \times 5$ $= (\overline{X} - \overline{Y}) \pm 1.96 \times 5$ = (20 - 10) ± 1.96 × 5 = 10 ± 9,8	20 ± Z

S4	a = 10 b = 25 (10-25$\pm$1.96	

L: Let's see Q6. It is in Chapter 9.

Q6. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the content is 500c.c. and write down your steps. α = 0.05

5 minutes later, L: Ok, the next question is very similar to this one. Basically there is only one condition different.

Q7. It is said by a drink seller that the content of their drink is over 500c.c.. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the drink seller is telling the truth, and write down your steps. α = 0.05

5 minutes later, L: I think you haven't prepared yet. We can stop here. I wonder if you would come for the third time after the final exam, when you really will have studied this part.

S3 and S4: Should be ok.

L: See you.

S3 and S4: Bye.

	Q6	**Q7**
Correct answer	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ = 500** H_1 **: μ ≠ 500 (two-tailed testing)** $Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$ **-** $Z_{\frac{\alpha}{2}}$ **= -1.96** **CR = {Z > 1.96 or Z < -1.96}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ $\geq$ 500** H_1 **: μ < 500 (left-tailed testing)** **-Zα = -Z0.05 = -1.645** **CR = {Z < -1.645}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$

	$= \frac{498.2 - 500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 > -1.96** **Accept H_0, the content is 500c.c.**	**= -1.8 / 1** **= -1.8 < -1.645** **Reject H_0, the drink seller is not telling the truth.**
S3	H_0 **:** μ ≥ 500 n = 100 **α = 0.05** H_1 **:** μ < 500 σ = 10 P ($\overline{X}$ > 498.2 / μ ≥ 500) = P ($\overline{X}$ - μ > 498.2 - 500) = P ($\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} > \frac{-0.8}{10 / \sqrt{100}}$) = P (Z > -0.8) = P (Z < 0.8) = 0.7881 Z= 1.96	
S4		

Q8. Which approach do you use to solve question 6 and 7? Why do you use this approach?

[I did not ask Q8 since S3 and S4 had not studied this part yet.]

Q9. Can you draw graph in question 6 and 7?

[I did not ask Q9 this time because I decided to observe whether S3 and S4 would actively draw graphs in the third interview after the final exam.]

6.14. The second student interview with S5 on 5th Jun 2009.

[In the second student interview, I asked S1&S2, S3&S4, S5 and S6-S10 to solve/answer nine questions. The interviews took place two or three weeks before the final exam, and many of them had not studied yet, especially the chapter of hypothesis testing. I tested Q6 and Q7 again in the third interview with S1-S5 whereas S6-S10 could not participate after the final exam finished.]

L: Hello, S5, thanks for coming today. This time I prepared nine questions for you, I would like you to solve them. The first question is testing how you understand the table of normal distribution. Please start solving the 5 sub-questions.

Q1. P (Z > a) = 0.67; P (Z > 0.67) = b; P (Z < c) = 0.67; P (Z < 0.67) = d and P (Z = 0.67) = e.

What are a, b, c, d and e?

S5 used table of type A and wrote ‘a’ = 0.43 and ‘b’ = 0.7486 (both wrong) very quickly. However, when he searched for ‘c’, he found something, compared tables of type A and type B, and browsed his notebook. Then he deleted the wrong answer for ‘b’, used type B and wrote down the correct answer for ‘b’. S5 used type A and found the answer(s) for ‘c’ [one correct and one wrong] and ‘d’. In the end, he used type B and wrote down the second answer for ‘a’ [still wrong].

10 minutes later, S5 finished.

Q1	**P (Z > a) = 0.67**	**P (Z > 0.67) = b**	**P (Z < c) = 0.67**	**P (Z < 0.67) = d**	**P (Z = 0.67) = e**
Correct answer	**a = -0.44**	**b = 0.2514**	**c = 0.44**	**d = 0.7486**	**e = 0**
S5 (type B)	P (Z > a) = 0.67 a = ~~0.43~~ 0.2514	P (Z > 0.67) = ~~0.7486~~ 0.2514	P (Z < c) = 0.44 or 0.45	P (Z < 0.67) = 0.7486	0.7486 – 0.2514

L: Let’s discuss your answers. In the first sub-question, you deleted 0.43, can you tell me how you found it?

S5: I found it from the table of smaller than (type A).

S5 pointed at 0.6664 corresponding to 0.43 which is just next to 0.67 on the second page of type A: I haven’t used this table for a long time, so I regarded bigger than [i.e. > in the question] as smaller than [which fitted the format of type A).

L: Are you pointing at 0.6664?

S5: Yes.

L: And what’s its neighbour?

S5: The same.

L: 0.67?

S5: Yes, it’s between 3 and 4.

L: Do you mean between 0.43 and 0.44?

S5: Yes.

L: Ok, then the second sub-question. Why did you answer 0.7486 in the beginning?

S5: It's the same [i.e. the same reason].

L: Can you show me where you found it?

S5 pointed at 0.7486 corresponding to 0.67.

L: Then how do you find that it's wrong?

S5: I looked at them and accidentally found that type B is bigger than and type A is smaller than.

L: So you found the answer of b 0.2514, can you point it out?

S5 pointed at 0.2514 corresponding to 0.67 correctly.

L: Ok, thanks, how about the 3rd sub-question? You should write P (Z < c) = 0.67, and c = 0.44 or 0.45, rather than P (Z < c) = 0.44 or 0.45.

S5: I see.

L: That's fine, where did you find the answer?

S5: That is smaller than.

S5 turned to the first page of Table A [negative z scores], and soon moved to the second page of Table A [positive z scores] and pointed at 0.6700 and 0.6736: And it's probably between these two.

L: 0.67?

S5: I mean 0.4 and 0.5, and both are 0.67, so I have two answers.

[Again, S5 did not think 0.67 = 0.6700. He believed that 0.67 is between 0.6700 and 0.6736.]

L: I know what you mean, and the fourth sub-question?

S5 used the second page of type A and found the answer for 'd' quickly.

L: Good, and the fifth one, is it bigger than by minus smaller than? [i.e. P(Z > 0.67) – P(Z < 0.67)]

S5: No, I wrote smaller than minus bigger than.

L: Why?

S5: Because I see the answer is still a *positive* value.

L: Ok, I see. Many of your answers are correct, but 'c' is only 0.44 because it just corresponds to 0.67; 'e' is 0, that I mentioned in the first interview, perhaps you forgot it.

S5: ?

L: Probability of Z equals any value would be zero.

S5: Infinite?

L: No, zero.

S5: I see.

L: Your answer for ‘a’ is not correct. P (Z > a) = 0.67, and ‘a’ cannot be found directly in type A and type B because there is no probability 0.67 in type B, and type A is about smaller than.

S5: …

L: P (Z > a) = 0.67, so what is P (Z < a)?

S5 pointed at his answer: Do you mean 0.2514?

L: No, 0.67 in the question means the probability which is in the table, but you use table of type B and find 0.67 corresponding to 0.2514 which is also a probability value. 0.67 is inside but you looked at the outside frame.

S5: Oh.

L: Let me ask you again, P (Z > a) = 0.67, so what is P (Z < a)?

S5: -0.67?

L: Can you write it down?

S5 considered for a while and wrote P (Z < a) = 0.5: is it 0.5?

L: Why?

S5 browsed his notes and pointed at P (Z < 0) = 0.5 in his note: I thought it is this, P (Z < a).

L: It’s P (Z < 0)

S5 deleted 0.5 and wrote -0.67: Oh, then it is -0.67.

L: But it becomes a negative probability

S5: You are right, so is it 1 – 0.67?

L: And what is it?

S5: 0.33.

L: Then what is ‘a’?

S5 used Table A. He did not see 0.33 on the first page, so he turned to the other page. Soon he came back to the first page: -0.42?

L: Where is it?

S5 pointed at .3372 corresponding to -.4 and .02: Here.

S5 found its neighbour: No, it’s -0.43.

L: What’s its neighbour?

S5: It is 0.33, so should be between these two.

L: You mean between -0.44 and -0.43?

S5: Yes.

L: Ok. Let’s see the next question.

Q2. Get two sets of independent random sample ($x_1, x_2, ... x_{20}$) & ($y_1, y_2, ... y_{20}$) from two normal populations X ~ N (20, 20^2) and Y ~ N (10, 10^2) respectively. The sampling distribution of ($\overline{X} - \overline{Y}$) follows N (a, b). What are a and b?

L: I designed this question according to an example in the textbook, the content is in Section 8-6, and you can use your textbook and notes.

S5: Is $\overline{X}-\overline{Y}$ different to a and b?

L: No. $\overline{X}-\overline{Y}$ is a random variable which follows normal distribution and its mean is ‘a’, and its variance is ‘b’.

S5 used his textbook and wrote the answer, till he wrote a =10.

	Q2
Correct answer	$\mu_1 = 20, \sigma_1 = 20, n_1 = 20; \mu_2 = 10, \sigma_2 = 10, n_2 = 20$ $\overline{X}-\overline{Y} \sim \mathrm{N}\,(\mu_1 - \mu_2, \frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2})$ $= \mathrm{N}\,(20 - 10, \frac{20^2}{20} + \frac{10^2}{20})$ $= \mathrm{N}\,(10, 25)$ a = 10, b = 25
S5	$\cancel{\overline{X} = 20, \overline{Y} = 10}$ $n_1 = 20, n_2 = 10$ $\sigma_1{}^2 = 400, \sigma_2{}^2 = 100$ $\overline{X}-\overline{Y} \sim N(\mu_1 - \mu_2, \frac{400}{20} + \frac{100}{10})$ 20 + 10 = 30 ~~a = 10~~ $\overline{X}-\overline{Y} = 0$ [S5 retried] a = 20, b = 10

L: Why does ‘a’ equal 10? I did not mean that you are wrong, but wanted to know how you get the answer.

S5: $\frac{400}{20} + \frac{100}{10}$ =30, and 30 minus ‘average of sample size’ to get 10.

L: Can you write it down?

[I just found that S5 had a misunderstanding of one word in the Chinese question. I wrote: sample sizes ‘both’ are 20. In Chinese, this word can also be interpreted as ‘mean’. Thus, S5 regarded $n_1 = 20$, $n_2 = 20$ as $\overline{X} = 20$, $\overline{Y} = 20$.]

L: Oh, I just found that my words in the question may cause your misunderstanding. Here I mean $n_1 = 20, n_2 = 20$, rather than $\overline{X} = 20$, $\overline{Y} = 20$.

S5: Then $\overline{X}-\overline{Y}$ could be 0, and a =20 and b = 10

L: What?

S5: This ($\frac{400}{20}$) is a, and this ($\frac{100}{10}$) is b.

L: Then what is $\mu_1 - \mu_2$?

S5: It is $\overline{X} - \overline{Y}$.

L: Equals 0?

S5: Yes. Isn't it?

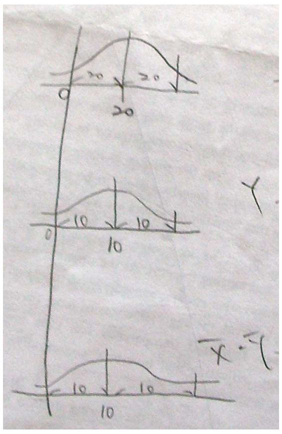
L: No. a = $\mu_1 - \mu_2$ = $\overline{X} - \overline{Y}$ = 20 – 10 = 10, and that's why I asked you why you wrote a = 10 in the beginning; you should remember that sample mean equals population mean. Also, $n_1 = 20, n_2 = 20$, and n represents the sample size, so b = $\frac{20^2}{20} + \frac{10^2}{20}$ = 25, not 30.

L: Now we know that X ~ N (20, 20^2), Y ~ N (10, 10^2) and $\overline{X} - \overline{Y}$ ~ N $(10, 5^2)$. If I draw the graph of X on top, can you draw other graphs below?

Q3. Following Q2, X ~ N (20, 20^2) is shown on top, please draw and Y ~ N (10, 10^2) and $\overline{X} - \overline{Y}$ ~ N (10, 5^2) in the second and third places.

L: In the first graph, the central 20 is the mean of X, and 20 in both sides represent the standard deviation.

[S5 drew the graph of Y which is the same size and shape to my graph of X, and marked 10s as its mean and standard deviation; then he drew the graph of $\overline{X} - \overline{Y}$, and also marked 10 as its mean and standard deviation.]

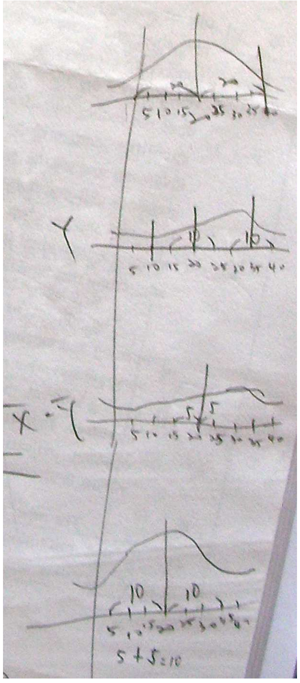


L pointed at the third graph: why do you write 10 here? The standard deviation is 5.

S5: …

L: Let me draw another graph of X with graduation on the x axis, can you draw graph of Y and graph of $\overline{X} - \overline{Y}$ again on such a scale?

[S5 drew the graph of Y and the graph of $\overline{X} - \overline{Y}$, but he felt that the graph of Y was not correct so he redrew it below the graph of $\overline{X} - \overline{Y}$, so the four graphs are X, abandoned Y, $\overline{X} - \overline{Y}$ and Y respectively).



L: Why are all the centres on 20? In the last graph, what does the left 10 mean?

S5 pointed at first 10 of Y ~ N (10, 10^2) in the condition: This.

L: How about the right 10?

S5: The same.

L: Then what does the centre 20 mean?

S5: I think, it is 5 away from 5 to 10, and also 5 away from 15 to 20, and 5 plus 5 equals 10.

L: Then how far away is from 5 to 20?

S5: From 20 to 25 or from 20 to 5?

L: Never mind. How do you deal with the sector from 10 to 15? Didn't you see it, or did you just ignore it?

S5: … I just added the two sectors…

L pointed at the first 10 of Y ~ N (10, 10^2): Ok. How do you deal with the 10?

S5: 5 plus 5 equals 10.

L: How does your 10 come?

S5: I think 10 is the distance. If I calculate and get the distance is 10, then I decide the points.

[S5 developed his own theories to solve the questions]

L: Ok, let's see the next question

Q4. Following Q2, what is the 95% confidence interval of $\overline{X} - \overline{Y}$?

S5: Isn't 95% confidence interval $Z_{0.025}$ **?**

L: You just write it down.

S5: And it equals 1.96.

L: Then?

S5: Its confidence interval equals 1.96, isn't it?

L: No.

S5 used his textbook and solved Q4.

	Q4	Q5
Correct answer	$\frac{1-0.95}{2}=\frac{0.05}{2}=0.025$ $Z_{0.025}=1.96$ $(\overline{X}-\overline{Y}) \pm Z_{0.025}\times\sqrt{\frac{400}{20}+\frac{100}{20}}$ **= (20 - 10) ± 1.96 × 5 =10 ± 9.8 = (0.2, 19.8)**	**They are significantly different because 0 is not included in (0.2, 19.8)**
S5	$Z_{0.025}=1.96$ $\overline{X}-\overline{Y} \pm Z_{0.025}\times\sqrt{\frac{400}{20}+\frac{100}{20}}$ (20-10) =10 ± 1.96 × 5	significantly different

L: Ok, it is correct. Do you remember what teacher A mentioned in class?

Q5. Following Q2, are the two populations significantly different?

S5: They are significantly different.

L: Why?

S5: If $\overline{X}=\overline{Y}$, $\overline{X}-\overline{Y}$ equals 0, then it is not significantly different.

L: You mean $\overline{X}-\overline{Y}$ has to be zero and then $\overline{X}$ and $\overline{Y}$ are not significantly different?

S5: Isn't it? $\overline{X}-\overline{Y}$ =10, so they are significantly different.

L: In fact, $\overline{X}$ and $\overline{Y}$ are significantly different, because its confidence interval does not include 0, and both ends are positive.

S5: Oh.

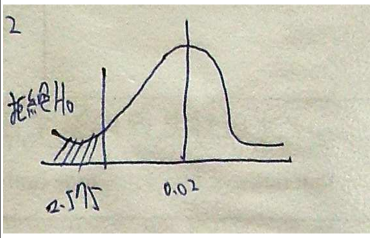
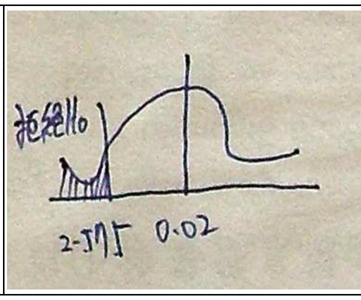
L: Let's see the next questions. They are in chapter 9.

Q6. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the content is 500c.c. and write down your steps. α = 0.05

Q7. It is said by a drink seller that the content of their drink is over 500c.c.. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the drink seller is telling the truth, and

write down your steps. α = 0.05

	Q6	**Q7**
Correct answer	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ = 500** H_1 **: μ ≠ 500 (two-tailed testing)** $Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$ **-** $Z_{\frac{\alpha}{2}}$ **= -1.96** **CR = {Z > 1.96 or Z < -1.96}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 > -1.96** **Accept** H_0**, the content is 500c.c.**	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ ≥ 500** H_1 **: μ < 500 (left-tailed testing)** $-Z_{\alpha} = -Z_{0.05} = -1.645$ **CR = {Z < -1.645}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 < -1.645** **Reject** H_0**, the drink seller is not telling the truth.**
S5	H_0 : μ = 500 H_1 : μ < 500 $Z = \dfrac{498.2-500}{\frac{10}{100}}$ $= \dfrac{-0.02\times 10}{100}$ $= \dfrac{0.2}{100}$ $= \dfrac{20}{1000}$ $= 0.02$ CR = {Z > 2.575} 0.02 < 2.575 reject H_0	H_0 : μ ≥ 500 H_1 : μ < 500 $Z = \dfrac{498.2-500}{\frac{10}{100}}$ = 0.02 CR = {Z > 2.575} 0.02 < 2.575 Reject H_0

L: Do you think that there is difference between Q6 and Q7?

S5: The condition in Q6 is ‘equal 500’, and the condition in Q7 is ‘bigger than 500’.

L: Are the testing processes the same?

S5: … should be the same.

Q8. Which approach do you use to solve question 6 and 7? Why do you use this approach?

L: I can see that you used the test statistic approach which is normally taught by teacher A.

S5: Yes.

L: You don’t use another method p-value approach.

S5: I can also use it.

Q9. Can you draw a graph in questions 6 and 7?

[Being asked, S5 drew two graphs but both of them were left-tailed without critical value and centre value.]

L: What are the critical ranges?

S5: 2.575.

L: And where will 0.02 be located?

S5: In the middle.

L: Then why do you reject both of them?

S5: …

L: You wrote that you reject H_0 because 0.02 is smaller than 2.575, but in your graph, 2.575 is in the left side of 0.02. What you draw is different to what you wrote.

S5: In my calculation, 0.02 is smaller than 2.575. ‘Smaller than’ is the same to H_1, so I have to reject H_0.

L: You mean that 0.02 is smaller than 2.575, the same to ‘smaller than’ in H_1?

S5: Yes.

L: Ok, thanks. In Q7, Z = $\dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$, but you missed the square root. Also, 498.2 – 500 = -1.8, not -0.2, so Z = -1.8. Besides, the critical range is $-Z_{0.05}$, can you find it in the table of normal distribution?

S5 pointed at .0537 and .0526: Between these two.

L: Can you find any closer one?

S5: Closer to what?

L: Closer to 0.05.

S5: .0505 and .0516.

L: I find that you often find the number whose first digits are the same or similar to what you are looking for as the answer, such as .0505 and .0516. However, .0495 and .0505 are the same close to .05, but you ignored it. In truth, the critical range is -1.645, between -1.64 and -1.65. It is negative, and you should compare -1.8 and -1.645. In Q6, it is two-tailed, so there are two critical ranges. They are $\pm Z_{0.025}$.

S5: 1.96.

L: Yes. I also noticed your other wrong calculation in Q6. The negative symbol just disappeared and 0.2/100 is not 20/1000.

S5: Oh, I calculated too quickly.

L: That’s fine, thank you so much today, see you soon.

S5: See you.

6.15. The second student interview with S6-S10 on 8th Jun 2009.

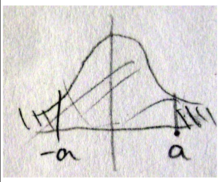
[In the second student interview, I asked S1&S2, S3&S4, S5 and S6-S10 to solve/answer nine questions. The interviews took place two or three weeks before the final exam, and many of them had not studied yet, especially the chapter of hypothesis testing. I tested Q6 and Q7 again in the third interview with S1-S5 whereas S6-S10 could not participate after the final exam finished.]

L: Hello, S6-S10, thanks for coming today. This time I prepared nine questions for you, I would like you to solve them. The first question is testing how you understand the table of normal distribution. Please start solving the five sub-questions.

Q1. P (Z > a) = 0.67; P (Z > 0.67) = b; P (Z < c) = 0.67; P (Z < 0.67) = d and P (Z = 0.67) = e.

What are a, b, c, d and e?

[Thirteen minutes later, S6 and S7 finished, but S8, S9 and S10 were still looking for the answer, so I told them to do Q2 if Q1 is done. Another 7 minutes later, I let them to stop solving and began asking them to explain how they found the answer for a, b, c, d, and e.]

Q1	**P (Z > a) = 0.67**	**P (Z > 0.67) = b**	**P (Z < c) = 0.67**	**P (Z < 0.67) = d**	**P (Z = 0.67) = e**
Correct answer	**a = -0.44**	**b = 0.2514**	**c = 0.44**	**d = 0.7486**	**e = 0**
S6 (type A & B)	P(Z > a) = 0.67 => 1 – P(Z<a) = 0.67 => P (Z<a) = 0.67 a = -0.44  	P (Z > 0.67) = b => b = 0.2514	P (Z < c) = 0.67 c = 0.44	P (Z < 0.67) = d = 0.7486 = d	P (Z = 0.67) = e = 1
S7 (type A & B)	P(Z > a) = 0.67 **a** => ~~0.44~~ ~~1.49 < a ≤ 1.50~~ 1.49 > a ≥ 1.50	P (Z > 0.67) = 0.2514	P (Z < c) = 0.67 c = 0.44	P (Z < 0.67) = d d = 0.7486	P (Z = 0.67) = e e = 0
S8 (type A & B)	1 – 0.67 = 0.33 a = ~~0.04~~ 0.404	0.2514	P(Z < 0.33) = ~~0.6255~~ 0.6293	0.7486	0

S9 (type A & B)		0.2514		-0.2514	0
S10 (type A& B)	-0.44	0.2514	0.44	0.7486	0

[I asked S9 first, because he did not answer any question, and I preferred to obtain his own explanation, rather than the words following other students who are more capable.]

L: Hi S9, can you point out where your answer for ‘b’ is in the table(s)?

S9 pointed at type B: P (Z > 0.67) = b, so ‘b’ = 0.2514.

L: Ok, you did not answer ‘c’, how about ‘d’?

S9 pointed at the first page of type A: P (Z < 0.67) = d, find 0.67, and find negative 0.67, that’s it.

L: So your answer is -0.2514?

S9: Should be. I am not sure.

L: Don’t you know there is another page?

S9: So? … Then the answer is 0.7486.

L: Ok. Did you *forget the existence of the second page of Type A*?

S9: I rarely use type A. Normally I use type B.

L: Ok, thanks.

L: S10, it’s your turn. Where is your answer for ‘a’?

S10 pointed at .6700 corresponding to .4 in the left and .04 on top on the second page of type A: Here.

L: So what is ‘a’?

S10: Because the symbol is reversed [i.e. > and <], so I add a negative symbol, and get -0.44.

L: Ok, ‘b’?

S10: I use type B, and directly find 0.67

L: How about ‘c’?

S10: I use the second page of type A, and directly find 0.67 in the table, then *find* the answer *out*.

L: Ok, then ‘d’?

S9: I find .6 in the left and .07 on top, and find the answer.

L: Ok, ‘e’?

S9: I guess.

S6 laughed.

L: Ok, thanks.

L: S8, how do you find ‘a’?

S8: I write 1 – 0.67 and get 0.33, and find 0.33 in the table of type B.

L: So what is ‘a’?

S8: 0.04.

L: Can you point it out?

S8 pointed at .3300 corresponding to .4 and .04 in the table of type B: This.

L: That is zero point?

S8: 0.404 (zero point four zero four).

L: Ok, write it down, and how about ‘b’?

S8: I directly find 0.67 in the table of type B and get 0.2514.

L: Ok, then ‘c’?

S8: I also write 1 – 0.67 and get 0.33, and find the answer in the table of type A.

L: Where is your answer?

S8 pointed at 0.6255: Here. Wait, perhaps I was wrong. It is 0.6293.

[S8 should find 0.67 within the table, but she found 0.33 in the frame.]

L: Ok, then ‘d’?

S8: Directly find it.

L: Ok, and ‘e’?

S8: 0.

L: For any reason?

S8: Because it is *equal*, so I think it is 0.

L: Thank you.

L: Hi, S7, please.

S7 pointed at *.0681* corresponding to 1.4 and .09 and .0668 corresponding to 1.5 and .00 in the table of type B: For ‘a’, 0.67 is between 0.681 and 0.668, but I cannot find 0.67, so I think that ‘a’ is between 1.5 and 1.49.

L: How about your previous answer, 0.44?

S7: I think it involves bigger than and smaller than, so I deleted it.

L: Ok, ‘b’?

S7: This is bigger than, so I directly use type B and get 0.2514.

L: Ok, ‘c’?

S7: It is smaller than, so I directly find 0.67 in the table of type A, and c is 0.44; ‘d’ is also smaller than, so I can directly use the table of type A and find ‘d’ = 0.7486.

L: Ok, then?

S7: I guess ‘e’ = 0, because ‘equal’ is not in the tables.

L: Ok, thanks.

L: S6, you are the last one.

S6: ‘A’ is bigger than or equalling to, but it cannot be found in the table of type B, so I have to use smaller than, and then add a negative symbol in front of the answer.

L: Why did you write 1 – P (Z < a) = 0.67?

S6 pointed at the graph he drew: Because 0.67 is the area of Z bigger than ‘a’, and it equals to 1 – the area of Z smaller than ‘a’ But this is about bigger than, and I use the table of smaller than, so I find 0.44 and I add a negative symbol to get -0.44.

[In S6’s graph, the shadow area is a smaller part, and he did not notice that 0.67 is bigger than 0.5.]

L: Ok, next one.

S6: I directly find the answer 0.2514.

L: Ok, ‘c’?

S6: It is smaller than, so I find c in the table of type B.

L: And ‘d’?

S6: Just find it, 0.7486.

L: ‘E’?

S6: I may have such an illusion that ‘equalling’ is 0.

L: Initially your answer was 1.

S6: Yes, but when I think about how to explain it, I feel that it is not 1.

L: Ok, thanks.

L: I designed the questions by myself, and I put the most difficult question ‘a’ in the beginning. I know that your teacher normally uses type B, but when the probability is bigger than 0.5 (i.e. half), you cannot find it in the table. Thus, I designed this question. In fact, you can transfer it to P (Z < a). Since P (Z > a) = 0.67, and what is P (Z < a)?

S8: 0.33.

S6: 0.33.

L: Yes, and you can find 0.33 in the table of type A and get ‘a’ = -0.44, right? All of you find ‘b’ = 0.2514. Comparing to ‘a’, ‘c’ =0.44 is easy because it can be found directly in the table. The answer of ‘d’ is 0.7486, do you notice that ‘b’ = P (Z > 0.67) = 0.2514, and ‘d’ = P (Z < 0.67) = 0.7486 and b + d = 1? Also, P (Z < c) = 0.67 and ‘c’ is 0.44, you should know that P (Z > a) = 0.67 and ‘a’ is -0.44 if you understand the *symmetry*. S6 drew a graph for ‘a’ and he got the correct answer. Although his graph is not totally correct, because 0.67 is bigger than a half and the shadow area should be a big part.

[S6-S10 did not respond.]

L: How many of you answered Q2? Let me see… only S6 and S7, that’s fine.

Q2. Get two sets of independent random sample ($x_1, x_2, ... x_{20}$) & ($y_1, y_2, ... y_{20}$) from

two normal populations X ~ N (20, 20^2) and Y ~ N (10, 10^2) respectively. The sampling distribution of ($\bar{X} - \bar{Y}$) follows N (a, b). What are a and b?

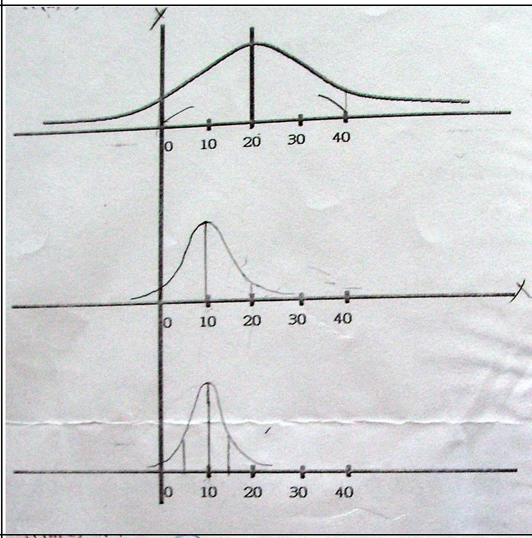
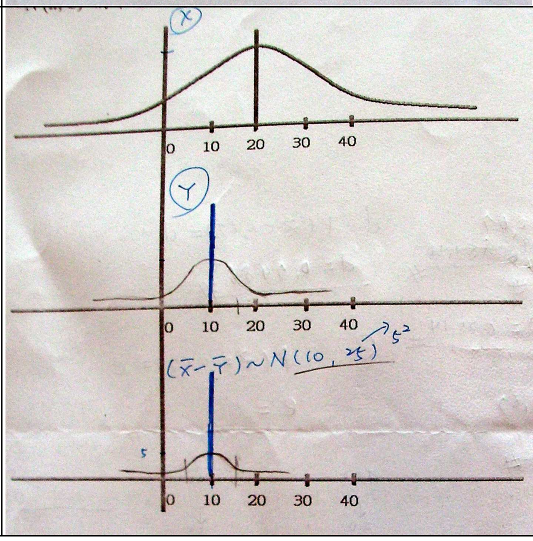
	Q2
Correct answer	$\mu_1 = 20, \sigma_1 = 20, n_1 = 20; \mu_2 = 10, \sigma_2 = 10, n_2 = 20$ $\bar{X} - \bar{Y}$ **~ N (** $\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ **)** **= N (20 – 10,** $\frac{20^2}{20} + \frac{10^2}{20}$**)** **= N (10, 25)** **a = 10, b = 25**
S6	$\mu_1 = 20, \sigma_1 = 20; \mu_2 = 10, \sigma_2 = 10; n_1 = n_2 = 20$ $\bar{X} - \bar{Y}$ ~ N (20 – 10, $\frac{400}{20} + \frac{100}{20}$) => $\bar{X} - \bar{Y}$ ~ N (10, 25) a = 10, b = 25
S7	$n = 20$ $\sigma_1^2 = 400, \sigma_1 = 20$ $\sigma_2^2 = 100, \sigma_1 = 10$ $\bar{X} - \bar{Y}$ ~ N (20 – 10, $\frac{400}{20} + \frac{100}{20}$) $\bar{X} - \bar{Y}$ ~ N (10, 25) a = 10, b = 25
S8	x Y
S9	
S10	

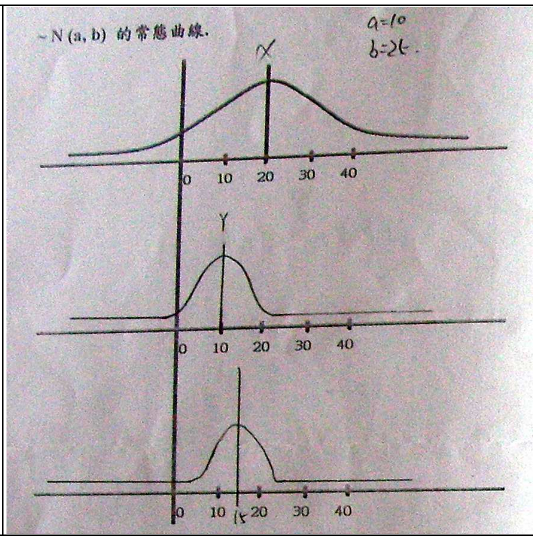
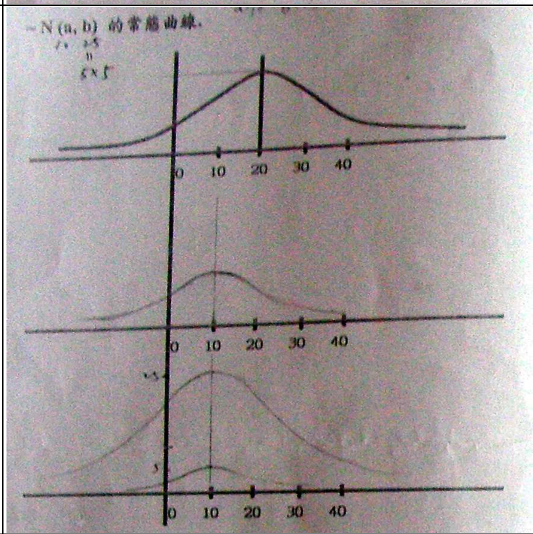
L: Q2 is on the first page of Section 8-6. I should tell you that the formula $\bar{X} - \bar{Y}$ ~ N ($\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$) on p. 279 is not correct. In fact, square root of should be removed. The correct format is on p.280, $\bar{X} - \bar{Y}$ ~ N ($\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$). We put variance or square of standard deviation after mean, rather than standard deviation. Thus, ‘a’ = 10, and ‘b’ = 25 or 5^2.

Q3. Following Q2, X ~ N (20, 20^2) is shown on top, please draw and Y ~ N (10, 10^2) and $\overline{X} - \overline{Y}$ ~ N (10, 5^2) in the second and third places.

[As printed, their original question showed that $\overline{X} - \overline{Y}$ ~ N (a, b), and I told them that a = 10, b = 25. However, I found that S6 and S9 saw 25 as standard deviation, so I reminded them that 25 is 5^2, in case they misunderstood the condition.]

10 minutes later, they finished.

	Q3
S6	
S7	

S8	
S9	
S10	

S10 did not draw.

L: Why didn’t you draw any graph?

S10: I don’t know how to draw, and I hate graphs.

L: That’s fine.

L: S8, why is there such a vertical line?

S8: Then we can clearly know that 20 is the top point.

[I intended to ask S8 why she drew the vertical line located at 10 in the graph of Y, but she thought I was asking about the vertical line in the first graph of X.]

L: How about the second graph?

S8: 10 is the top point of Y.

L: So the top point is Y?

S8: No, it’s 10. Y = 10.

L: Mean of Y is 10?

S8: So I draw 10 as the top point.

L: Ok, and the third graph?

S8: I don’t know. I just subtract the two calculated means.

L: What?

S8: 25 – 10.

L: So 15, ok. Why is the second graph narrower than the first one? Why does the curve of Y look different to the curve of X?

S8: Different?

L: Are your second graph and third graph in the same shape but different location?

S8: Yes.

L: Don’t you feel that the first curve is wider than yours?

S8: But there is a difference between 10 and 20.

[S8 thought that the spread of normal curve is relevant to its mean, but later she had a different explanation.]

L: Did you tend to draw a curve of Y in the same width as the curve of X?

S8: Yes.

L: But why do they become narrow?

S8: Oh, is that what you mean?

L: Yes.

S8: Hmm… it is sharper. I didn’t consider this.

L: Do you just consider where its centre is?

S8: I think it is because that 10 is a smaller number.

L: 10 is smaller than 20, so it is narrower.

S8: Yes.

L: I see, thanks.

L: S9, you drew a vertical line located at 10 in the second graph? What does 10 mean?

S9 pointed at the first 10 in the question: $Y \sim N(10, 10^2)$: Here.

S9 pointed at the horizontal line he drew: And here is a square, 20 [he was pointing at the mean of X, as the width of the square] and 20 [he was pointing the standard deviation of X, as the height of the square]. Comparing to this, here is also a square, 10 [the mean of Y, as the width of another square] and 10 [the standard deviation of Y, as the height of another square].

L: You see it as a square, so the curve becomes smaller.

S9: Yes.

L: For the third graph, did you draw the higher curve initially?

[S9 drew two curves of $\overline{X} - \overline{Y}$.]

S9: Yes. I thought the height is 25, but you said it is 5^2, so I redrew it. Also, a is 10.

L: So it becomes flatter.

S9: Yes.

[He thought standard deviation influences the height of the curve.]

L: Ok, thanks.

L: S6, why is the centre located at 10?

S6 pointed at the condition: I thought that 10 [i.e. mean of Y] is the x-coordinate and 10 [i.e. standard deviation of Y] is the y-coordinate, so I drew like this.

L: Why is the curve of Y narrower than the curve of X?

S6: Because it is 10^2, and the first one is 20^2.

L: Why is 20^2 wider?

S6: 20 is bigger, so the range is bigger,

L: How about the third graph? You drew a wider curve in the beginning.

S6: Yes, I thought it is 25^2, so I drew a very wide curve, but you reminded us it is 5^2, not 25^2, so I drew a narrower curve.

L: Thanks.

L: S7, please.

S7: I feel that in the first graph, the x-ordinate is 20, and the height seems corresponding to 20. So in the second graph, the centre is 10, and I think that its height should be also 10.

L: Where do you get the y-coordinate?

S7 pointed at 10^2 of Y ~ N (10, 10^2) in the condition: Here.

L: How about the third curve?

S7 pointed at 10 of $\overline{X} - \overline{Y} \sim$ N $(10, 5^2)$: The centre is 10.

He pointed at 5^2 of $\overline{X} - \overline{Y} \sim$ N $(10, 5^2)$ and continued: And it is 5^2, so the height is only 5.

L: Ok, thank you. Does any one have other ideas?

S6-10: …

L: Ok. Let me explain. For the first graph, the first 20 is mean, and the second 20 is standard deviation. Do you remember what empirical rules are about?

S6-S10: …

[Their textbook mentioned the contents of empirical rules, but not the name, so the students looked confused when I asked.]

L: I don't know if teacher A taught that to you or not, but the content is that 68% of area is between the interval of mean plus and minus one standard deviation, 95% of area is between the interval of mean plus and minus two standard deviations, and 99% of area is between the interval of mean plus and minus three standard deviations. Thus, the second curve is narrower than the first one, and the third curve is even narrower.

L: The centres of second and third curves both locate at 10.

S8: Why 10?

L: It is decided by the mean, a = 10. Not 25 – 10. Thus, the second curve is narrower than the first one, and the third curve is even narrower because of their standard deviation.

S7: How about the height?

L: The sizes of the two populations are not mentioned, so we cannot know how tall they are, but the samples from two populations are 20, so the height of the third curve should be lower.

S7: Can you tell me where it is wrong?

L: Basically your curves are correct, but the third curve should be thinner than the second one, because its standard deviation is smaller than the second one.

S7: I see.

L: Ok, let's see the next question.

Q4. Following Q2, what is the 95% confidence interval of $\overline{X} - \overline{Y}$?

L: It is in 8-6. Q5 is also taught by teacher A, so you can answer it when you finish Q4.

Q5. Following Q2, are the two populations significantly different?

Z	Q4	Q5
Correct answer	$\frac{1-0.95}{2} = \frac{0.05}{2} = 0.025$ $Z_{0.025} = 1.96$ $(\overline{X} - \overline{Y}) \pm Z_{0.025} \times \sqrt{\frac{400}{20} + \frac{100}{20}}$ **= (20 – 10) ± 1.96 × 5 =10 ± 9.8 = (0.2, 19.8)**	**They are significantly different because 0 is not included in (0.2, 19.8).**
S6	$1-\alpha = 0.95$ $\alpha = 0.05$ confidence interval = (20 - 10) ± 1.96 $\times \sqrt{25}$ = 10 ± 1.96 $\times 5$ = 10 ± 9.8 = (0.2, 19.8)	The interval does not include 0, so the two populations are significantly different.
S7	95% => 1.96 95% confidence interval ⇨ [(20 - 10) ± 1.96 $\times \sqrt{\frac{20^2}{20} + \frac{10^2}{20}}$] = (10 ± 1.96×5) = (10 ± 9.8) = (0.2, 19.8)	The interval does not include 0, so there is a significant difference.

S8	$\sqrt{5}$	
S9		
S10		

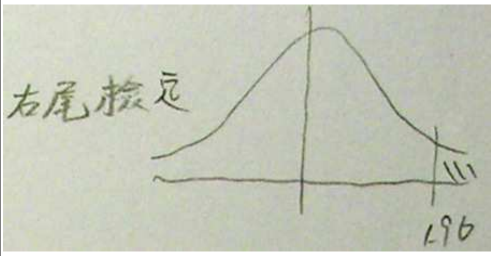
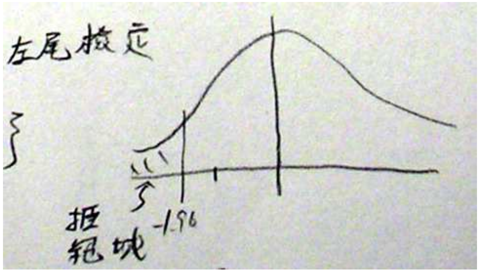
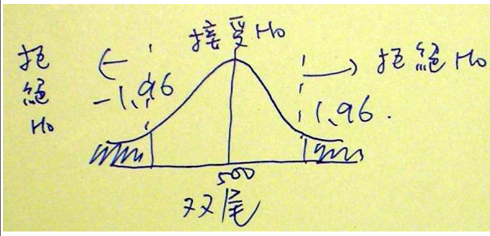
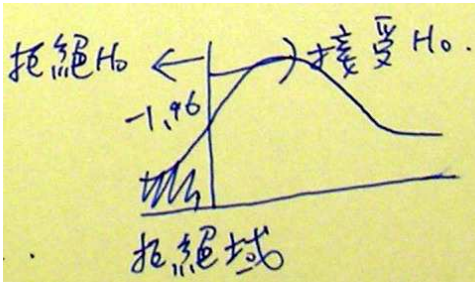
L: Ok, only S6 and S7 have answers. You can find the formula in the textbook and the confidence interval is $(\overline{X} - \overline{Y}) \underline{+} Z_{0.025} \times \sqrt{\frac{400}{20} + \frac{100}{20}}$. Both of you are correct.

L: Now we only have 10 minutes left, so please do Q6 and Q7 together. I suggest you compare the conditions first.

Q6. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the content is 500c.c. and write down your steps. α = 0.05

Q7. It is said by a drink seller that the content of their drink is over 500c.c.. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the drink seller is telling the truth, and write down your steps. α = 0.05

	Q6	**Q7**
Correct answer	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ = 500** H_1 **: μ ≠ 500 (two-tailed testing)** $Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$ **-** $Z_{\frac{\alpha}{2}}$ **= -1.96** **CR = {Z > 1.96 or Z < -1.96}** $Z_0 = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $= \frac{498.2 - 500}{\frac{10}{\sqrt{100}}}$	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ ≥ 500** H_1 **: μ < 500 (left-tailed testing)** $-Z_{\alpha} = -Z_{0.05} = -1.645$ **CR = {Z < -1.645}** $Z_0 = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $= \frac{498.2 - 500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 < -1.645**

	= -1.8 / 1 **= -1.8 > -1.96** **Accept H_0, the content is 500c.c.**	**Reject H_0, the drink seller is not telling the truth.**
S6	$H_0: \mu \leq 500$ $H_1: \mu > 500$ right-tailed testing $\alpha = 0.05$ $Z_{0.025} = 1.96$ CR = {Z > $Z_{0.025}$}  $Z_0 = \dfrac{498.2-500}{10/\sqrt{100}}$ $= \dfrac{-1.8}{1}$ = -1.8 $\ngtr$ -1.96	**H_0 : μ ≥ 500 H_1 : μ < 500** left-tailed testing CR = {Z < - $Z_{0.025}$}  $Z_0 = \dfrac{498.2-500}{10/\sqrt{100}}$ = -1.8 < -1.96 Accept H_0 [S6 did not write down the decision until I asked him to.]
S7	$H_0: \mu = 500$ $H_1: \mu < 500$ or $H_1: \mu \neq 500$ two-tailed testing  (because H_0 is true, the result is to 'not reject H_0') judgment correct, it is true accept H_0	$H_0: \mu \geq 500$ $H_1: \mu < 500$ left-tailed testing  H_0 is fake, the result is to reject H_0 judgment correct, it is true
S8	$\sigma = 10, \alpha = 0.05, Z_{0.05} = 1.645,$ $n = 100, \overline{X} = 498.2$ $H_0: \mu = 500$ $H_1: \mu < 500$ CR = {Z > $Z_{0.05}$ = 1.645}	$H_0: \mu \geq 500$ $H_1: \mu < 500$
S9	$\mu_0 = 500cc, n = 100, \overline{x} = 498.2cc,$ $s = 10cc, \alpha = 0.05$ $H_0: \mu \geq 500$ $H_1: \mu < 500$	$\mu_0 > 500cc, n = 100, \overline{x} = 498.2cc,$ $s = 10cc, \alpha = 0.05$ $H_0: \mu \geq 500$ $H_1: \mu < 500$

	CR = {T < $t_{0.05}(99)$ = 1.671 ~ 1.658}	CR = {T < $t_{0.05}(99)$ = 1.671 ~ 1.658}
S10	H_0 : μ > 500 H_1 : μ < 500 $Z_{0.05}$ = 1.645	

L: Ok, we don't have much time, let's discuss your answer now. S10, you did not write much, did you study this part yet?

S10: No.

L: That's alright. In Q6, most of you, including previous five participants, make H_0 as μ > 500, and H_1 as μ < 500. I designed Q7 to let you have a chance to compare the conditions. In fact, H_0 : μ = 50, H_1 : μ ≠ 500. Some of you have the correct null hypothesis but the wrong H_1 : μ < 500, and still do the one-tailed testing. The opposite of '=' is '≠', and Q6 is a two-tailed testing. Q7 is one-tailed testing.

L: S7, can you explain how you decide to reject or accept H_0 in Q6?

S7: See 500 as the middle, 498.2 should be in the accept region.

L: How do you know?

S7: I guessed.

S8 laughed.

L: Didn't you calculate?

S7: How can I calculate without numbers?

L: There is the formula in the textbook, you can read it later.

S7: Ok.

L: S6, you did not finish Q7, can you write down to accept or reject H_0 ?

S6: Accept.

L: Where is -1.8?

S6 made a mark in the right-hand side of -1.96: -1.8 is here.

[S6 wrote -1.8 < -1.96 in his solution of Q7.]

L: Ok, thanks. We have finished now, thank you so much. See you tomorrow.

S6-S10: bye.

Q8. Which approach do you use to solve question 6 and 7? Why do you use this approach?

[They had to go to class, so I did not have time to ask. In Q6 and/or Q7, S6 compared Z_0 and CR; S8 found CR as a z score and S9 found CR as a t score; S7 found z score $\pm$1.96 and S10 found a z score 1.645. Their solutions indicated that they all used test statistic approach.]

Q9. Can you draw a graph in questions 6 and 7?

[S6 and S7 actively drew graphs after writing hypotheses.]

[I invited S6-S10 to participate in the third interview after they finished the final examination, but some of them had part-time jobs, and some of them were home-sick, so they did not come.]

6.16. Preparation for the second teacher interview.

In the first student interview, ten participants were asked to look for probability in the tables of normal distribution.

P (Z < -0.03) = 0.488: seven were correct, three wrote 0.512.

P (Z < 1.22) = 0.8888: six were correct, three wrote 0.1112.

P (Z > -3.1) = 0.999: six were correct, three wrote 0.001.

P (Z > 2.05) = 0.0202: seven were correct, others wrote 0.202, 0.0207 and 0.9798.

P (-1.4 < Z < 2.32) = 0.909: four were correct, two did not answer, four were wrong.

P (-2.41 < Z < -0.03) = 0.48: two were correct.

P (Z = 3) = 0: three were correct.

We can see that

1. P (a < Z < b) was more difficult for the participants. Less than half of them correctly answered them.
2. Many of them did not know/notice that P (Z = something) = 0.
3. In the first four questions, many mistakes might be relevant to *symmetry*. Many students did not notice or consider the relationship between the area and the z score. (For instance, the area > 2.05 should be a small part less than 0.5).

The ten participants also were asked to solve two questions by using the table of t distribution.

t(0.005, 9) = 3.25: eight were correct, one did not answer and one thought d. f. = 9 – 1 = 8.

t(0.99, 23) = -2.5: only one was correct. Many students could not find 0.99 in the table of t distribution and did not apply the idea of *symmetry*.

In the second student interview, finding 'a' in P (Z > a) = 0.67 is difficult for students. Only 2 students find the correct answer -0.44, and three students write 0.44. Ignoring *symmetry* is still a common problem for students.

Q3 showed that seven students knew that the centre of normal curve is located in its mean, and five students knew that standard deviation influences the spread.

Six students made correct hypotheses in Q7, but only one student made correct hypotheses in Q6. Most of them did not correctly interpret the questions. Many participants were asked by the teacher to memorise $z_{0.05} = 1.645, z_{0.025} = 1.96,$ and $z_{0.005} = 2.575$, but they often used wrong ones, such as $Z_{0.05}$ = 2.575. Some participants use the critical value $z_{\frac{\alpha}{2}}$ even if they made hypotheses for one-tail testing. One student regarded Z_0 = -1.8 < critical value -1.96. Another student gave a positive critical value even if he interpreted it as a left-tail testing. He wrote Z_0 = 0.02 < critical value 2.575 and rejected H_0. In his graph, 0.02 was in the accept region and critical value 2.575 was on the left-hand side of 0.02.

6.17. The second teacher interview on 11th Jun 2009.

L: Hello, teacher A. This is the summary of my interview with students, please have a look.

[3 minutes later.]

TA: We can see that they cannot solve probability which involves interval.

L: Some students can solve it, but most of them cannot. Besides, they feel difficult when it involves positive and negative.

TA: This can be improved by using graphs, and I did draw graphs in class.

L: In the first student interview, the questions all ask P (Z > or < some value) = ? Some students are correct but some others have a reverse answer (such as giving the probability of P (Z > a) as the answer of P (Z < a) or vice versa).

TA: They lack the concept of *direction*.

L: Yes, and many of them often mix up negative and positive Z score. In the interviews, I asked them to explain how they find the answers step by step. They use table of type A and table of type B.

TA: Left-tailed and right-tailed.

L: Yes. The table of type A covers negative and positive Z scores, but some students only look at the first page. They seem not to notice the existence of the other page. Also, P (Z > -3.1) cannot be found directly in both tables.

TA: It can be transferred to P (Z < 3.1).

L: Yes, but some students don't have the idea of *symmetry*.

TA: In fact, it is because of their distraction in class, and I did teach these ideas. Also, their fundamental ability in mathematics is not good.

L: Right, they made many mistakes in calculation. The first time I noticed the weak mathematics ability of students was five years ago when one student wrote $\frac{1}{2}+\frac{1}{3}=\frac{2}{5}$. I was really shocked.

TA laughed.

L: In my research, my focus is not on their mathematics ability but how they solve statistics questions and what difficulties they may have. For example, I am interested in the influence of using table of type A and type B. In the first interview with S1 and S2, I only gave S1 the table of type B. He had some trouble in using type B because he preferred type A. He did not ask me to give him table of type A, so he was wrong in some questions. I asked him why he did not ask, and he said he thought I was testing how they realise table of type B. I let S2 use my textbook, so he could use his preferred table.

TA: They don't care about '>' or '<'.

L: Yes, one of them said that he (or she, I forgot which student) can *only* use the table in such a discipline, and cannot deal with other kind of questions. I know that you draw graphs in class, but not many students draw. Some students ignored the negative symbol, such as seeing P (Z > -3.1) as P (Z > 3.1).They may also mix up ‘>’ and ‘<’. However, when facing P (Z < - something) …

TA: They know to use symmetry. I stressed that in class.

L: And P (Z > -something) is 1- P (Z > something).

TA: Yes.

L: But they cannot find the answer if they directly use the table. In fact, asking ‘>’ or ‘<’ doesn’t make a difference, but asking the probability of an interval is harder for them. I did not ask the probability of an interval between two Z scores because the other two questions may be transferred to this.

L: I also find that they don’t understand the relationship between the area and ‘>’, ‘<’, ‘+’ and ‘-’.

TA: I see.

L: Eight students correctly answered the first question of t-distribution. S7 thought the degree of freedom as number size, so she subtracted 1 from the degree of freedom in order to get the degree of freedom. Only one of them correctly solved the second question because the answer cannot be found in the table of t-distribution directly. There is only the right-tailed table of t-distribution, and the probabilities in the table are rather small. The only correct student took advantage of ‘symmetry’.

[TA became very asleep because he just finished his upgrade examination two days before the interview. He did not talk much.]

L: In the second interview, question 1 asks 5 sub-questions all involves 0.67, but this time I gave them the probability and let them find the Z score out. The first sub-question (P (Z > a) = 0.67, find ‘a’) is more difficult because it cannot be found directly in both tables. In table of type B, there is no answer; and it is necessary to transfer to P (Z < a) = 0.33 in order to use table of type A. Thus, only 2 students have the correct answer -0.44. Many students answer 0.44, because in the table of type A, 0.67 corresponds to 0.44, but they don’t notice the table is about ‘<’.

TA: Hmmm.

L: One student (S5) very frequently found neighbour(s) of correct answer as his answer, such as finding 0.6736 corresponding 0.43, and writes 0.43 as the answer; or finding 0.0256 corresponding to 1.95 when looking for 0.025 corresponding to 1.96. He made such a mistake in 3 of the 6 questions. Another (S8) interprets 0.67 corresponding to .4 in the left and .04 on top as 0.404. One student interprets .0202 as 0.202.

TA: Will they have an advanced course of statistics?

L: I am not sure, I know that they will have an optional course of statistics software, such as SPSS, and the statistical knowledge may be applied in their project in the last year.

L: In the 3rd question in the 2nd interview, I drew the curve of $X \sim N(20, 20^2)$ as example, and asked them to draw the curve of $Y \sim N(10, 10^2)$ and the curve of $\overline{X} - \overline{Y} \sim N(10, 25)$. Many students know the centre of the curve is decided by the mean, but one thought the mean is 25 – 10.

TA: I didn't teach them this part.

L: This part should be taught in the first semester when teacher C was teaching. The students say that they didn't learn this. However, some students interpret 25 as the standard deviation, rather than the variance or square of standard deviation. Also, 2 students think that the height of the curve is influenced by the standard deviation.

TA: I don't know how to draw it.

L: The centres of both curves locate at 10.

TA: How about the standard deviation? And how can we decide the width and height of curve?

L: According to the *empirical rules*, 68% of area will be in the range of mean plus and minus 1 standard deviation.

TA: Right.

L: That means normal curve, even if it is not standard normal curve, should be in the same/similar shape. I mean, if the standard deviation is smaller, the curve should be flatter.

TA: If the standard deviation is smaller, the curve should be sharper, because it is more centralised.

L: What I mean is comparing 2 *normal* curves, the curve with smaller standard deviation should be flatter, since they are normal, and the percentage of area in the range of mean plus and minus some standard deviation is *constant*. In fact, I still have to confirm my thought. Comparing X and Y, Y has a smaller standard deviation, so it is more centralised, but since X and Y are both normal, I think their curves should be in the same shape, Y is narrower, and it is also shorter. I wonder, if we draw the curve of Y in a sharper shape, is the area between 1 standard deviation still 68%? And will this curve still be a normal curve?

TA: …

L: Some students thought the mean 20 and standard deviation 20 decide the centre and the height of the curve. I feel this is an issue, because they don't know the influence of standard deviation to the 'spread'. When I asked them to explain, they said in the graph of X, there is a square.

TA: How about the hypothesis testing?

L: In Q7 in the second student interview, six students made correct hypotheses. However, in Q6, only one student correctly made two-tailed testing hypotheses, others made one-tail hypotheses and some students made right-tailed hypotheses. In the solving process, they use $Z_{0.05}$, many students memorise $Z_{0.05}$ = 1.645, but some students such as S5 write $Z_{0.05}$ = 2.575 which equals to $Z_{0.005}$. S7 know that Q6 is two-tailed testing and Q7 is left-tailed testing, she mistakenly memorises $Z_{0.05}$ = 1.96 ($Z_{0.05}$ = 1.645) and uses it as the critical value in both questions. She did not use $\pm Z_{\frac{0.05}{2}}$ as the critical value in Q6. S6 thought Q6 is right-tailed and Q7 is left-tailed testing, but his critical value in both questions is not $Z_{0.05}$, but $Z_{\frac{0.05}{2}}$. I remember in the last week, S6 solved a similar question in the blackboard, and he wrote 1.96, rather than correct 1.645.

TA: They do not memorise correctly.

L: I wonder if they mix this up with the confidence interval.

L: S5 made a series of mistakes in Q6. For instance, he wrote critical region is {Z > 2.575}, but in his graph, 2.575 is in the negative side, and the shadow reject area is in the left-hand side of 2.575. His Z_0 is 0.02, and he wrote 0.02 < 2.575, so reject H_0, but in his graph, 0.02 is in the right-hand side of 2.575. His ideas are chaotic, and what he draws seems independent to what he writes. I will test him again after the final.

TA: He doesn’t link the knowledge. Sorry I don’t help much, I almost fell asleep.

L: I still should say thank you. Let’s have a lunch together.

TA: Ok, let’s go.

6.18. Questions of the third student interview.

Part I

Q1. In the final examination, do you solve the questions by yourself, or memorise the answers given by the teacher?

Q2. After the one-year statistics course, how do you feel about the course? Do you have any personal feelings or question about it?

Q3. What is your feeling about the teacher's statistics classes?

Q4: What aspect do you suggest the teacher improve?

Q5. How do you think about the practice questions in class?

Part II

Q6. How do you understand statistics?

Q7. What do you think statistics can do?

Q8. Do you like statistics?

Part III

Q9. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10c.c. Please test whether the content is 500c.c. and write down your steps. $\alpha = 0.05$

Q10. It is said by a drink seller that the content of their drink is over 500c.c.. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10c.c. Please test whether the drink seller is telling the truth, and write down your steps. $\alpha = 0.05$.

6.19. Third student interview with S1-S5 on 25th Jun 2009.

[In the second student interview which took place two or three weeks before the final exam, some participants had not studied the chapter of hypothesis testing yet, and S1-S5 agreed to participate in the third interview after the final exam. I tested Q6 and Q7 in the second interview again as Q9 and Q10 in the third interview without prior notice. I also ask them some other questions.]

L: Hello, S1-S5, thanks for coming today. In this interview, I will ask you a few questions, please answer in order from S1, S2 … to S5.

Q1. In the final examination, do you solve the questions by yourself, or memorise the answers given by teacher A?

S1: I solve them by myself.

L: Since the answers were given, why don’t you just memorise them?

S1: They are so easy that I don’t have to memorise them.

L: Very easy?

S1: Yes.

L: Do you feel any question is difficult?

S1: Perhaps the last question. I am not sure if it is correct or not, because teacher A did not give us the answer for this question.

L: How do you know if you are correct or not in other questions?

S1: Before the final, I taught some classmates several times.

L: Then you may have an impression of the answers. I found that some numbers are not easily calculated.

S1: Right.

L: Do you think this may cause your miscalculation?

S1: No.

L: Ok, thanks.

L: S2, do you solve them by yourself?

S2: No, I memorise them.

L: Why don’t you solve them by yourself?

S2: Because I don’t know how, and they are difficult.

L: Which part do you find difficult?

S2: The *formulas*. There are plenty of formulas, and I don’t remember what the *symbols* represent.

L: Symbols also make you confused?

S2 and S4: Yes.

L: Ok, thanks.

L: S3?

S3: I solve them by myself because I don’t like to memorise the answers.

L: Do you feel any question in the final exam is difficult?

S3: The last one, because its numbers are larger and I may enter a wrong number when using the computer.

L: When getting the mean and variance?

S3: Yes.

L: How about other questions?

S3: No problem.

L: How about formulas?

S3: No problem.

L: Ok, S4.

S4: I memorise a half part of the answers and solve the other half by myself. I write the formulas, and try to calculate to get answers close to the answers given by teacher A. I don’t solve the last question.

L: You mean you memorise the formulas, but calculate by yourself.

S4: Yes.

L: Which question is the most difficult?

S4: The last one.

L: Thanks.

L: S5.

S5: When I practise with the questions, I repeat solving them by myself, and I find that I memorise the processes and numbers naturally.

L: Do you feel any question is difficult?

S5: The last one. I stop about half-way.

L: Why?

S5: Because I feel my answer is strange, and I cannot get the result.

L: Ok, thanks.

Q2. After the one-year statistics course, how do you feel about the course? Do you have any personal feeling or question about it?

S1: I feel if you already know something, then you can calculate to get other things. Unlike calculus, most of us never learned statistics before. Calculus is taught in many high schools, especially in the industry area. Their theories are similar, but formulas in statistics are more complicated. Some classmates feel statistics is very difficult, but I feel it is not so difficult.

L: Where do you feel it is difficult?

S1: Using wrong formulas. I made this mistake before.

L: How do you know?

S1: I asked teacher A, and he told me.

L: You mean once teacher A asked you to solve a question in chapter 9 on the blackboard?

S1: Yes, the 32^{nd} question.

L: It is about hypothesis testing of population mean, and what formula did you use?

S1: I used the formula of hypothesis testing of population ration p.

L; Ok, you can add any comment later if you have some.

S1: Ok.

L: S2.

S2: I don’t know.

L: Any idea or feelings?

S2: Statistics is very complex.

L: Compared to what?

S2: I feel it is as difficult as mechanics.

L: I don’t like physics, too.

S2: I hate to see the symbols and numbers.

L: Do you feel the symbols are difficult?

S2: No. I cannot find a word to describe my feeling. I must remember many formulas, and I cannot solve the questions if I do not remember the formulas.

L: If I design a question for you, will you be able to find the formula according to the conditions?

S2: No. But I may be able to solve it if I can use the book.

L: Thanks.

L: S3.

S3: I feel that statistics is more flexible than calculus, because calculus just requires putting the numbers into the formula.

L: Why is statistics different?

S3: Statistics involves testing and software. It is not limited to paper work. I think it is more diversified.

L: Ok, thanks

L: S4.

S4: Statistics is more difficult.

L: Compared to what?

S4: Calculus.
L: Why?
S4: It is deeper than calculus.
L: In which aspect?
S4: Formulas and questions.
L: Can you explain?
S4: …
L: Do you mean that many questions in calculus just require you to calculate and get the answer, and many questions in statistics are words describing the condition and you have to find the correct formula by yourself?
S4: Yes.
L: Thanks.

L: S5.
S5: I feel statistics is difficult, and also easy.
L: Why?
S5: There are too many formulas, and I get confused about the data of population and of sample.
L: Can you give an example in the textbook?
S5: It may take a long time to find it.
L: That’s alright.
S5: For example, the question mentions ‘all students’, and I think it means the population, but it is the sample.
S1: I think S5 means the 5th question in the final exam.
S5: Let me have a look, yes, it is. I think this question wants us to find sample data.
L: Do you mean why use the result of sampling to infer population?
S5: No. I often don’t know the difference between population and sample.
L: What do you think is their difference?
S5: One is about the whole, and the other is about sampling, but I still made mistakes in calculation.
L: What do you mean, mistakes in calculation?
S5: The question looks like it is about population, but the calculated result is about sample. *To deal with statistics questions, you have to distinguish population and sample*.
S1-S4 laughed.
S5: Do you know what I mean?
L: I am thinking.
S1: I don’t know.
S5: Sometimes I regard large distribution as small distribution, or small distribution as

large distribution [S5 meant sample size bigger or smaller than 30]. For example, I don't know whether the fifth question is small or large distribution.

L: What do you think?

S5: Small [i.e. sample size].

L: Why?

S5: Because there are two departments.

L: Wow, it is really the best sentence of today.

S5: The question asks about a large sample, but I think it is a small sample because there are only two departments.

L: In fact, you should decide the formula according to the condition. Can you explain this, S1?

S1: This is a large sample, because they sample 64 persons in each department, and it is bigger than 30.

L: You are right, n is 64, not 2.

S5: I always get confused about this. Besides, I sometimes get confused about estimation of population ration p and estimation of difference of two population means $\mu_1 - \mu_2$. I cannot decide which it is according to the question. The formulas are really too many, so I cannot remember all of them and often get confused in problem-solving. However, statistics is also easy because you can get the result if you find and use the correct formulas. It is rather theoretical.

L: For me, the difficulty of statistics is to decide the type of question and corresponding formulas according to the conditions.

S5: Yes.

L: On p. 380 and p. 381 in your textbook, there are flows of mean testing and variance testing showing conditions and their corresponding situation and formulas. Some conditions are not taught in class.

S5: That's what I mean by 'too many formulas'. Sometimes I *remember some formulas* but I *forget what its condition is*.

L: In my previous class, I once gave students the preparatory questions before the final exam, and changed the order of the questions in the final exam. Most students couldn't recognise what situations the questions were and simply copied the formulas and methods from the preparatory sheet.

Q3. What is your feeling about teacher A's statistics classes?

S1: Once I told teacher A to speed up.

L: Is he too slow?

S1: He repeated explaining to help everyone understand. I think 1 or 2 times are enough, but he may repeat the same type of question for four or five times.

L: Thanks.

S2: He would repeat teaching a question till each student understands.
L: Do you understand?
S2: I was not listening.
L: So he did not teach for enough times?
S2: No. It depends on whether you listen to him or not.
L: Why don't you listen to him?
S2: Because S1 sets me up.
L: Thanks.

S3: I feel teacher A uses different metaphors, such as 'a bad pear pretends to be an apple' and 'seeing diamond as stone' to explain type I error and type II error. I think I will remember them forever. His descriptions are closer to life, unlike the textbook.
L: It makes you feel impressed.
S3: Right.
L: Thanks.

S4: I like the atmosphere in his class.
L: He really talks a lot.
S4: I feel relaxed. He talks and does teach something.
L: He doesn't tell you any jokes in class.
S1: But he himself is funny.
L: Thank you.

S5: At the beginning of classes, he reviews the previous content, and then he uses examples to introduce the new formulas and contents, and calculates to get the results. What I feel the most impressive is that he usually says, 'S5, you study so hard, I love you so much, please come to solve the question on the blackboard.' Also, he usually gets mixed up and writes the wrong answer on the blackboard.
L: I also notice that.
S5: His teaching style is special. Normally he *explains in his words* first, rather than following the steps in the textbook. It makes it easier to understand. For example, when he taught sampling distribution of sample ratio, he did not waste time on the genteel words in the textbook, but wrote the abstract and key words of this area to help us build a clear understanding of what it is (video 7:22).
L: He explains first.
S5: Yes. I cannot understand what the textbook is talking about, so I feel teacher A's method is good.

L: Thanks.

Q4: What would you suggest for teacher A to improve?

S1: I think his teaching is quite simple for students in the middle level, because he helps you to build a good foundation and then your knowledge accumulates slowly in his teaching. However, many students have never learned this before and they are not good in mathematics. Although they study hard, they still cannot understand. On the other hand, some students do not study hard but they can understand this subject easily, it is so unfair.

L: Do you mean yourself?

S1 kept laughing: No. in a word, he tried to help less capable students learn better.

L: I see, thanks.

S2: He usually gives us examples that are easy to understand, and it helps us learn faster.

L: Why do you disappoint him?

S2: I am uninterested in mathematics.

L: How about statistics?

S2: I learn something that may be used in exams.

L: Thanks.

S3: I think he is a teacher so he knows what students are thinking. He uses some approaches closer to students to teach them.

L: What do you mean, approaches closer to students?

S3 does not explain.

S4: His explanation is easy to understand, he is close to students, and he uses simple approaches to teach us.

S5: He always writes something on the blackboard first. It helps us understand statistics better because I cannot understand what the textbook is talking about. After he explains in his words, I can realise what the textbook is about.

L: Do you have any suggestions to improve anything? This is my original question.

S1: No.

S2: This is good enough.

S3: Agree.

S4: He is good.

S5: I feel that he often makes a wrong calculation and writes the wrong answer, this is what he has to improve.

L: Teacher A only teaches you statistics in the 2nd semester, right?
S1: Yes, but he taught us calculus in the 1st year.
L: How about S6-S10?
S1: No, they were taught by another teacher. Teacher A is more familiar with us.

Q5. What do you think about the practice questions in class?

S1: They are aimed to increase our understanding, to test how to use the formulas, and to prevent the fear of examinations. They also train our braveness because teacher A wants us to solve them on the blackboard.
S2: They help us to know which formula should be used to solve each kind of question, and we know what formulas we should use in the examinations.
S3: I think they are different to the examples because there are not solutions of practise questions in the textbook, and we have to think by ourselves.
L: So you can really practise.
S4: He explains examples and lets us solve practice questions.
L: Do you practise?
S4: No.
S5: He always explains the examples, and lets us solve the practise questions. I think they are not too different. He never lets us practise questions that he did not teach or that are too difficult. He only lets us practise easy questions.
L: Thanks.

Q6. How do you understand statistics?

S1: Not much. Statistics is mysterious. We are just taught a part of it and I don't know what the later part is. Once my friend asked me what S.D. is, but I don't know how to answer him.
L: S.D. is standard deviation. Does anyone else know?
S1-S5: …
S1: I think, statistics is more than what I have learned. It can be deeper and more complicated, so I think I don't know it much.
S2: I don't know much, because teacher A only taught us the basic part of statistics for business. I think the advanced part should be learned in masters or doctoral level.
S3: I don't know much, too. It is not only about 'writing', but also 'seeing'. Some people know what it is if they see the conditions. For example, I may forget sample S.D. and population S.D. soon.
L: Me, too. I may forget some details occasionally.
S4: I don't understand. I *cannot distinguish the conditions* when reading the questions.
S5: I don't know the later part which teacher A did not teach. For the part we were taught, I only know how to use the formulas to find the answer. I don't know the

origins of the formulas, and even why there is statistics [i.e. the subject]. I only know how to calculate and find the answer. I also don't know how statistics can be used outside the classroom. It can be useful or useless in daily life.

L: As teacher A said, to prevent 'a bad pear pretends to be an apple'.

Q7. What do you think statistics can do?

S1: Statistics can be used to manage, test, measure, and…

S2: Testing.

S1: And some charts, such as the ticket accumulation of election, and life of a light bulb. It is really frequently used in life, so I think statistics is useful.

S2: In my daily life, I never met any situation to use statistics. Teacher C (in the 1^{st} semester) used PowerPoint in the class, rather than writing on the black board.

L: What is in his PowerPoint?

S2: The examples and solutions, the same as the textbook. But sometimes he added some graphs.

S3: I think it can be used in many circumstances, such as computer, design, and daily life.

S4: Statistics can be used in daily life, such as total numbers, percentage.

S5: I think statistics is useful, for example, to see if someone exceeds the speed limit. I think statistics can be used in those situations which require calculation, such as gambling and probability. We don't use it, but we know the process.

Q8. Do you like statistics?

S1: I like it because I learn some knowledge to use; but I also don't like it because of its complexity. Even if I was given the condition, I may still be unable to solve it.

L: You mean calculation?

S1: No. I mean you have to consider which formulas you should use according to the given conditions. The questions in a series are really hard.

S2: I like it because it may help me find a good job which requires many kinds of skill; I also don't like it because it may be totally useless in my entire life.

S3: I think I like it, because it can train our logic. Also, my memory of graphs is not good, and I can train my ability with graphs by drawing graphs by myself.

S4: I don't like it because it is very complicated, especially the conditions.

S5: I like it because it can be used in daily life; but I wish the data could be simplified.

L: In your second interview, some of you had not studied yet, so I would like you to solve the two hypotheses testing questions again.

Q9. It is shown that the content is 500c.c. on the package of a drink. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation

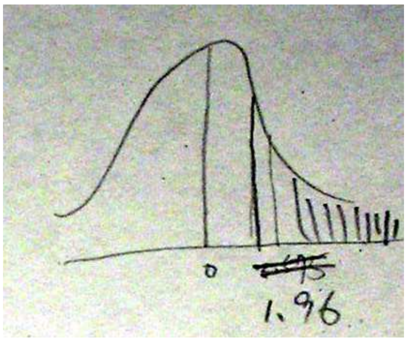
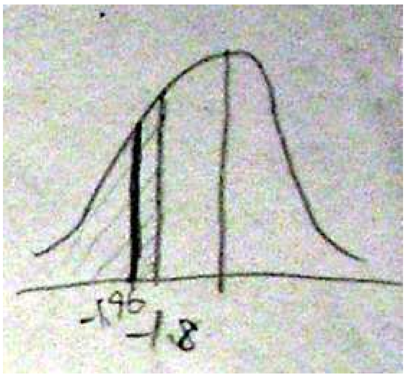
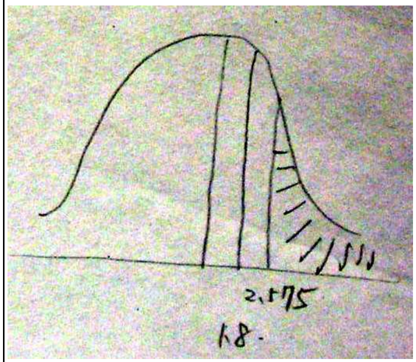
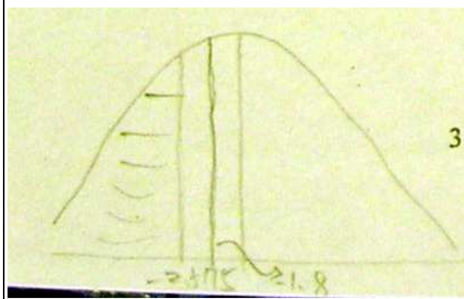
10 c.c. Please test whether the content is 500c.c. and write down your steps. α = 0.05

Q10. It is said by a drink seller that the content of their drink is over 500c.c.. Now an institute tests 100 bottles, and gets the average 498.2c.c. and standard deviation 10 c.c. Please test whether the drink seller is telling the truth, and write down your steps. α = 0.05

S1 murmured: 1.96
L: Do you remember that or do you find it in the textbook?
S1: I remember that, that is enough.
L: Please don't read it out.
S1: Sorry.
S1 finished in 7 minutes, and others stopped in about 20 minutes.

	Q9	**Q10**
Correct answer	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ = 500** H_1 **: μ ≠ 500 (two-tailed testing)** $Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$ **-** $Z_{\frac{\alpha}{2}}$ **= -1.96** **CR = {Z > 1.96 or Z < -1.96}** $Z_0 = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \frac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 > -1.96** **Accept** H_0**, the content is 500c.c.**	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ ≥ 500** H_1 **: μ < 500 (left-tailed testing)** $-Z_{\alpha} = -Z_{0.05} = -1.645$ **CR = {Z < -1.645}** $Z_0 = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \frac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 < -1.645** **Reject** H_0**, the drink seller is not telling the truth.**
S1	H_0 : μ = 500 H_1 : μ ≠ 500 (2 tails testing) 0.05 ÷ 2 = 0.025 [S5 added this	H_0 : μ > 500 H_1 : μ ≤ 500

	after he finished this question.] $\dfrac{\mu - \overline{X}}{\sigma / \sqrt{100}}$ $= \dfrac{498.2 - 500}{10 / \sqrt{100}}$ = -1.8 > -1.96 Accept H_0 It is true.	$\dfrac{\mu - \overline{X}}{\sigma / \sqrt{n}}$ $= \dfrac{498.2 - 500}{10 / \sqrt{100}}$ = -1.8 $\ngtr$ ~~-1.96~~ -1.645 Reject H_0 Not true.
S2	Z = $\dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{500 - 498.2}{\frac{10}{\sqrt{100}}}$ $= \dfrac{0.18}{1}$ = 0.18 It is true. [When I asked S2 to explain what $\overline{X}$ and μ mean, he wrote $\overline{X}$ above 500c.c. and μ above 498.2c.c; he used a correct formula and got a positive test statistic Z.]	Z = $\dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{500 - 498.2}{\frac{10}{\sqrt{100}}}$ [deleted] $= \dfrac{0.18}{1}$ [deleted] It is true.
S3	[On S3's question sheet, she reversely wrote $\overline{X}$ above 500c.c., and μ above 498.2c.c.; she used a correct formula and got a positive test statistic Z_0. S3 also wrote n above 100 bottles, σ above 10c.c. and α above 0.05.] H_0 : μ = 500 H_1 : μ < 500 $\dfrac{\alpha}{2} = 0.025$	[On S3's question sheet, she reversely wrote μ = 498.2. She used a correct formula and got a positive test statistic Z_0.] H_0 : μ > 500 H_1 : μ < 500 α = 0.05 $Z_0 = \dfrac{500 - 498.2}{10 / \sqrt{100}}$ = 1.8

	$Z_0 = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = \frac{500-498.2}{10/\sqrt{100}} =$ 1.8 90 $95 \rightarrow 0.05$ $99 \rightarrow 0.1$ $Z_{0.05} = \cancel{2.575}\ 1.96$  ~~It is true.~~ [Later she rewrote and got the correct Z_0] $Z_0 = \frac{498.2-500}{10/\sqrt{100}}$ = -1.8  It is true.	$Z_{0.05} = 2.575$  It is true. [Later she rewrote and got the correct Z_0.] Z = $\frac{498.2-500}{10/\sqrt{100}}$ = -1.8  It is true.
S4	$H_0: \mu \leq 100$ $H_1: \mu > 100$ $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = \frac{500-100}{\sqrt{498.2/10}}$ $= \frac{400}{\ } ?\, Z_{0.05} \rightarrow 1.645$	$H_0: \mu \geq 100$ $H_1: \mu < 100$ $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$
S5	~~n = 500~~ n=5 $\overline{X}$ = 498.2	

	$\sigma = 10$ H_0 : μ =498.2 H_1 : μ≠498.2 $Z_0 = \dfrac{498.2 - 500}{10 \big/ \sqrt{500}}$ [deleted] $Z_0 = \dfrac{498.2 - 500}{10 \big/ 10}$ $= \dfrac{-1.8 \times 10}{10} > 1.96$ = -1.8 < 1.96 Reject H_0, testing standard is not true.	H_0 : μ ≥ 500 H_1 : μ < 500 $Z_0 = \dfrac{498.2 - 500}{10 \big/ 10}$ $= \dfrac{-1.8 \times 10}{10} > 1.96$ = -1.8 < 1.96 Reject H_0, testing standard is not true.

L: I will ask how you solved them. S4 you wrote the least, so you are the first one to answer. Can you tell me how you decide the hypotheses?

S4: …

L: Why is there *100*?

S4: 100 bottles.

L pointed at μ: What is this?

S4: That is μ.

L: What does μ mean?

S4: Is it *sample number*?

L: No. I can see that you find the formula, but it seems you *don't know what each symbol means*.

L: S2, you found the correct formula. What is this [I pointed at $\overline{X}$]?

S2 wrote $\overline{X}$ above 500c.c.: this.

L: And what is this (I pointed at μ)?

S2 wrote μ above 498.2c.c.: this. σ is 10 and 100 is the number of bottles.

L: How do you get the result from 0.18?

S2: I thought *0.18 should compare to standard deviation*.

L: And?

S2: Smaller than standard deviation.

L: So?

S2: So it is true.

L: Who taught you that?

S2: No one.

L: How about Q10? Why did you delete it?

S2: I think Q10 is about ‘more than’, so it should be different to Q9.

L: How?

S2: I don’t know.

L: S5, why is μ = 498.2 in H_0 ?

S5: Because it says ‘gets the average 498.2c.c.’.

L: Why did you compare the negative result to 1.96?

S5: if -1.8 is $\boldsymbol{\mu}_1$, and 1.96 is $\boldsymbol{\mu}_2$ **.**

L: Why is there $\boldsymbol{\mu}_2$ **?**

S5 pointed at H_0 : μ = 498.2 : assume $\mu_1 = \mu_2$,

H_1 : $\mu \neq 498.2$: assume $\mu_1 \neq \mu_2$.

S5: Pointed at -1.8 < 1.96: because $\mu_1 \neq \mu_2$, so reject H_0 .

L: Where does 1.96 come from?

S5: Significant level.

L: Significant level 0.05, so you get 1.96?

S5: Yes.

L: How about Q10?

S5: The same, I also compare μ_1 and μ_2 .

L: Why do you reject H_0 because -1.8 < 1.96?

S5: $\mu_1 < \mu_2$ *(-1.8 < 1.96), so* $\mu_1 < \mu_2$ *(μ < 500).* I don’t know, maybe I am wrong.

L: S3, why do you have two sets of answers?

S3: Because I am *not sure about which one* $\overline{X}$ *is and which one* μ *is***.**

L: Which set do you think is correct?

S3: Should be the second set.

L: Why did you change your choice?

S3: I think one question should be ‘reject’.

S1 laughs.

L: Why?

S3: Normally teachers don’t give questions with the same answer.

L: You did not answer ‘true’ or ‘false’ in the second try of Q10, what do you think?

S3: Should be ‘true’

L: In Q10, do you accept or reject H_0 ?

S3: To accept H_0 **.** H_0 : μ > 500, so it is *one tail testing*. α = 0.05, so the critical value is 2.575, but it is negative because Z_0 is negative.

L: α = 0.05, it is …

S1: 1.96

S4 and S5: 1.96.

S3: Am I wrong?

L: I don’t memorise them, because it is easy to make a mistake.

S1: It is 1.645. It would be 1.96 if α is divided by 2.

L: You wrote 2.575, it corresponds to *0.005*.

L: S1 is correct in both Q9 and Q10. He has a clear thought of this, although he doesn’t draw the graphs. In S3’s Q9, there should be two shadow areas, but she only draws one. In her Q10, the critical value is -1.646, so -1.8 is in the reject range. Some students may compare -1.8 and -1.96 and see -1.8 < -1.96, but you don’t give such an answer.

S5: I do.

L: No, you compare -1.8 and +1.96. S1 compares -1.8 and -1.645 and he knows -1.8 < -1.645, so he has correct answer.

L: Our interviews are all finished. Here is your payment. Thank you so much for participating in the study.

S1-S5: You are welcome, thank you, too. Bye.

6.20. In vivo codes in the second and third student interviews.

(IN VIVO) CODES	OPEN CODES	CODES
Q1. **P (Z > a) = 0.67; P (Z > 0.67) = b; P (Z < c) = 0.67; and P (Z < 0.67) = d.**	Question format; > & <; Experience; Reverse thinking; Symmetry	28 25 37 38 39
In second sub-question, S2 moved his finger in the positive page of type A and stopped near P (Z < 0.45) = .6736 but soon moved away, and stopped at P (Z < 0.44) = .6700	Choice of table: S2 used tablea; > & <; Mis-interpreting	16 25 40
1 – 0.67 = 0.23	Mis-calculation	41
S2 pointed at 0.6293 with .3 in the left and .03 on top: this one.	Mis-interpreting	40
smaller than 0.67….. I don't know.	> & <; Symmetry	25 39
S2 did not use table B.	Choice of table	16
In Q1.1, S2 did not know > & <; in Q1.3, S2 saw probability 0.33 as z,	> & <; Misinterpreting	25 40
Bigger than,… smaller than,.. negative	> & <; Negative; Symmetry	25 27 39
S1 used both table A and table B	Choice of table	16
S2 used reverse approaches of using table A when solving Q1.1 & Q1.3	Mis-interpteting; Inside-out & outside-in	40 42
$\overline{X} - \overline{Y} \sim$ N ($\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$) in textbook	(In)correctness of material	43
L: wider because of smaller variance? S1: right, and its beginning part is flatter. It's not as sharp as second curve. L: its variance, 25, is smaller, so the curve is flatter. Is that what you mean? S1: yes.	Self theory: Variance smaller ⇨ Flatter ⇨ Looks wider	5 44 10
L: where is the part of square root? S2: am I wrong to directly multiple the answer in second question?	Mis-transcribing	45

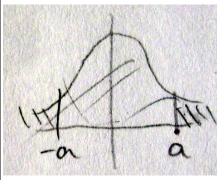
(19.8, 0.2)	Number line	15
The claim is real because the standard deviation is 10c.c. and the difference of 498.2 and 500 is only 0.2c.c. Thus 100 bottles should be within it.	Self theory; Miscalculation	5 41
S2 did not know hypothesis	Hypothesis	46
Many questions in your textbook only ask the principles of hypothesis testing, but I want you to finish the process.	Material	47
L: α equals 0.05 S1: confidence interval?	Confused of α and confidence interval	18 48
S1 drew a normal curve and a left tail, but he was not sure and murmured: this side or that side?	Which side	49
500 × 0.05 = 25 > 10; so reject. (500c.c. and under the level 0.05, so multiply them, and get 25. It's 15 bigger than standard deviation.)	Self theory	5
H1 is confirmed, the claim is true.	Conflict of H_0 and H_1	5, 11
The claim is true because if the standard deviations $\pm$ 10 add mean, then it will exceed 500	Self theory	5
H_0 = 500 H_1 < 490	Material; Teacher	47 50
S1 drew a graph when solving Q6, but the mean of his graphs are always NOT in the centre.	Self theory	7 5
S1 marked 500 between 498.2 and 508.2, but closer to 508.2	Number sense	15
P(Z > a) = 0.67, P(Z<-a)=1-0.67; P(Z > a) = 0.67, 1 - P(Z > a) = 0.67, P(Z > a) = 0.33	Mis-transfering; Conflict	51 11
b is the area, and 0.67 is α (0.67 should be z)	Symbol	48
S4 mix up the location of z score and probability	Format; Mis-interpreting	28 40
Cannot directly find it in the table	Not directly	33
S4 did not know that a = 10 is the centre and b = 25 represents degree of diverge	Definition	44 52
In S3's graphs, She did not notice the location of values on the number line	Unfamiliarity	7 15
Empirical rules		53
When finding 95% confidence interval of $\overline{X}-\overline{Y}$, S4 regarded a = 10 as $\overline{X}$ and b = 25 as $\overline{Y}$, and	Symbols; Mis-transcribing	48

ignored the square root of the formula		
S3 and S4 used test statistic approach in Q6 and Q7	Strategy	36
S5 compared table A and table B…	Compare and revise	54
S5 pointed at 0.6664 corresponding to 0.43 which is just next to 0.67 in the second page of table A	Mis-targeting	55
I haven't used this table for a long time	Unfamiliarity	30
L: what's its neighbour? S5: the same…. 0.67 is between 0.43 and 0.44	Between	15 35
I looked at them and accidentally found that table B is bigger than and table A is smaller than.	> & <	25 54
You should write P (Z < c) = 0.67, and c = 0.44 or 0.45, rather than P (Z < c) = 0.44 or 0.45	Writing; Unreasonable equal	28
0.4 and 0.5, both are 0.67, so I have two answers. [S5 did not think 0.67 = 0.6700. he thought 0.67 is between 0.6700 and 0.6736.]	Between; Number line	35 15
I wrote smaller than minus bigger than… because I see the answer is still a *positive* value.	Fundamental discipline	10
L: P (Z > a) = 0.67, so what is P (Z < a)? S5: do you mean 0.2514? (P (Z > 0.67) = 0.2514)	Mis-interpreting; Format	40 28
S5 had four answers here: P (Z > a) = 0.67 P (Z < a) = (1) 0.2514 (2) -0.67 (3) 0.5 (in S5's note, P (Z < 0) = 0.5) (4) 1 – 0.67 = 0.33	 (1) mis-interpreting; (2) symmetry; (3) mis-reading; (4) area	 40 39 56 6
P (Z < a) = 0.33 S5 pointed at .3372 corresponding to -.4 and .02: here. S5 found its neighbour: no, it's -0.43,… it is 0.33, so should be between -0.44 and -0.43.	Mis-targeting; Between	55 35
is $\overline{X}-\overline{Y}$ different to a and b?	Symbol	48
$\frac{400}{20}+\frac{100}{10}$=30, and 30 minus 'average of sample size' to get a = 10.	Self theory	5
S5 had a misunderstanding of one word in the chinese question. I wrote: sample sizes 'both' are 20. in chinese, this word can also be interpreted as	Mis-reading; Statistic language	56 57

‘mean’. thus, S5 regarded $n_1 = 20, n_2 = 20$ as $\overline{X}$ = 20, $\overline{Y}$ = 20.		
S5: then $\overline{X} - \overline{Y}$ could be 0, and a = $\frac{400}{20}$ = 20 and b = $\frac{100}{10}$ = 10 L: what is $\mu_1 - \mu_2$? S5: it is $\overline{X} - \overline{Y}$. L: equals 0? S5: yes. Isn’t it?	Self theory (perhaps caused by misunderstanding of symbols)	5 48
Centres of S5’s both graphs are 20 (S5 did not know that the center is decided by mean)	Prior knowledge	10
L: In the graph of Y, what does the left 10 mean? S5 pointed at first 10 of Y ~ N (10, 10^2) in the condition: this L: how about the right 10? S5: the same. L: then what does the centre 20 mean? S5: I think, it is 5 away from 5 to 10, and also 5 away from 15 to 20, and 5 plus 5 equals 10. L: then how far away is from 5 to 20? S5: from 20 to 25 or from 20 to 5? L: never mind. How do you deal with the sector from 10 to 15? Didn’t you see it, or did you just ignore it? S5: … I just added the two sectors… L pointed at the first 10 of Y ~ N (10, 10^2): ok. How do you deal with the 10? S5: 5 plus 5 equals 10. L: how does your 10 come? S5: I think 10 is the distance. If I calculate and get the distance is 10, then I decide the points.	Self theory	5
L: what is the 95% confidence interval of $\overline{X} - \overline{Y}$? S5: isn’t 95% confidence interval $Z_{0.025}$? ... and it equals 1.96. L: then? S5: its confidence interval equals 1.96, isn’t it?	Memory	13 30 5

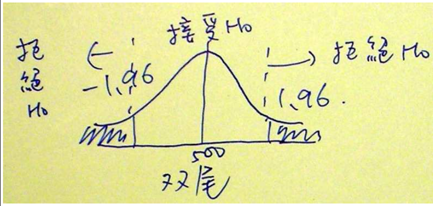
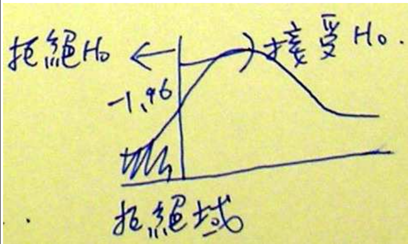
If $\overline{X} = \overline{Y}$, $\overline{X} - \overline{Y}$ equals 0, then it is not significantly different… Here $\overline{X} - \overline{Y}$ =10, so they are significantly different.	Self theory; Explained by words	5 57
$Z = \dfrac{498.2-500}{\frac{10}{100}} = \dfrac{-0.02\times 10}{100} = \dfrac{0.2}{100} = \dfrac{20}{1000} = 0.02$	Mis-transcribing (square root); Mis-calculation; Mis-transferring (negative symbol); Mis-calculation	45 41 51
S5: the condition in Q6 is 'equal 500', and the condition in Q7 is 'bigger than 500'. L: are the testing processes the same? S5: … should be the same.	Cannot distinguish conditions	57 58
S5 used test statistic approach in Q6 and Q7		36
L: What are the critical ranges? S5: 2.575.	Memory; Negative	13 27
L: and where will 0.02 be located? S5: in the middle.	Self theory	5
L: you wrote that you reject H_0 because 0.02 is smaller than 2.575, but in your graph, 2.575 is in the left side of 0.02. What you draw is different to what you wrote. S5: in my calculation, 0.02 is smaller than 2.575. 'Smaller than' is the same to H_1: $\mu < 500$, so I have to reject H_0.		7 11
L: the critical range is $-Z_{0.05}$, can you find it in the table of normal distribution? S5 pointed at .0537 and .0526: between these two. L: can you find any closer one? S5: closer to what? L: closer to 0.05. S5: .0505 and .0516. L: I find that you often find the number whose first digits are the same or similar to what you are looking for as the answer, such as .0505 and .0516. However, .0495 and .0505 are the same close to .05, but you ignored it. In truth, the critical range is -1.645, between -1.64 and -1.65. It is negative, and you should compare	Mis-targeting; Negative; Between; First digits	55 27 35 40, 59, 15 49

-1.8 and -1.645. In Q6, it is two-tailed, so there are two critical ranges. They are $\pm Z_{0.025}$.		
[Some students are confused and did not distinguish confidence interval and critical region.]		18
S9: P (Z < -0.67) = 0.2514, so P (Z < 0.67) = - 0.2514 L: Did you *forget the existence of second page of Type A*? S9: I rarely use type A. Normally I use type B.	Negative probability; Unknow the second page (positive) of table A; Unfamiliarity	30
P (Z > a) = 0.67 S10: P (Z < 0.44) = 0.67, so P (Z > -0.44) = 0.67. S10: because the symbol is reversed [i.e. > and <], so I add a negative symbol, and get -0.44.	Cannot directly find; > and <; Negative	33 25 27
I use the second page of table A, and directly find 0.67 in the table, then *find* the answer (c) *out*.	Inside out	42
I write 1 – 0.67 and get 0.33, and find 0.33 in the table of type B	Ignore influence of '1 –' (> should become <)	51
a is 0.04… a is 0.404	Mis-interpreting	40
P (Z < c) = 0.67 [available process => P (Z > c) = 0.33] S8: P (Z < 0.33) = 0.6255…0.6293	Mis-interpreting; Mis-targeting; Inside-out V.S. outside-in	40 55 42
S7 pointed at *.0681* corresponding to 1.4 and .09 and .0668 corresponding to 1.5 and .00 in the table of type B: for 'a', 0.67 is between 0.681 and 0.668, but I cannot find 0.67, so I think that 'a' is between 1.5 and 1.49. 1.49 > a $\geq$ 1.50	Mis-interpreting; Number line	40 15
It's bigger than, so I use table B…; it's smaller than, so I use table A (Find students' mechanism of choosing table and following process)	Choice of table	16
I guess e = 0 because 'equal' is not in the table	Guess; Equal; Not in the table	36 33
'a' is bigger than or equalling to, but it cannot be found in the table B, so I have to use smaller than, and then add a negative symbol in front of the	Cannot be found; > and <; Negative	33 25 27

answer. (‘>’ => ‘< -’)		
P(Z > a) = 0.67 => 1 – P(Z < a) = 0.67 => *P (Z < a) = 0.67 (S1 wrote this only for using the table)* a = -0.44 L: why did you write 1 – P (Z < a) = 0.67? S6 pointed at the graph he drew: because 0.67 is the area of Z bigger than ‘a’, and it equals to 1 – the area of Z smaller than ‘a’. [In S6’s graph, the shadow area is a smaller part, and he did not notice that 0.67 is bigger than 0.5.] But **this (question) is about bigger than, and I use the table of smaller than, so I find 0.44 and I add a negative symbol to get -0.44.** 	Compare area; Illogical formula (maybe just ignore because of mis-calculation); Unmatched graph; Symmetry; Transferring	6 11 41 7 60 39
I maybe have such an illusion that ‘equalling’ is 0	Memory	13
Comparing to ‘a’, ‘c’ =0.44 is easy because it can be found directly in the table	Directly found	33
I should tell you that the formula $\overline{X}-\overline{Y}$ ~ N $\left(\mu_1-\mu_2, \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$ in p. 279 is not correct.	Incorrectness of material	43
S6 and S9, interpreted 25 as standard deviation, so I reminded that 25 is 5^2, in case they misunderstood the condition.	Format; Variance and standard deviation.	28 44
I don’t know how to draw, and I hate graphs.	Hate graphs	8
L: S8, why is there such a vertical line? (I intended to ask her why she drew the vertical line located at 10 in the graph of y, but she thought I was asking the vertical line in the first graph of x). S8: then we can clearly know that 20 is the top point. L: how about the second graph? S8: 10 is the top point of Y. L: so the top point is Y?	Self theory	5

S8: no, it's 10. Y = 10. L: mean of Y is 10? S8: so I draw 10 as the top point. L: ok, and the third graph? S8: I don't know. I just subtract the two calculated means. L: what? S8: 25 – 10. (the centre of curve locates at 'variance – mean') L: so 15, ok. Why is the second graph narrower than the first one? Why does the curve of Y look different to the curve of X? S8: different? L: are your second graph and third graph in the same shape but different location? S8: yes. L: don't you feel that the first curve is wider than yours? S8: but there is a difference between 10 and 20 (S8 thought that (1) the spread of normal curve is relevant to its mean, but later she had different argument) L: did you tend to draw a curve of Y in the same width to the curve of X? S8: yes. L: but why do they become narrow? S8: oh, is that what you mean? L: yes. S8: hmm… it is sharper. (2) I didn't consider this. L: do you just consider where its centre is? S8: I think it is because that 10 is a smaller number. L: (3) 10 is smaller than 20, so it is narrower? S8: yes.		
L: S9, you draw a vertical line located at 10 in the second graph? What doe 10 mean? S9 pointed at the first 10 in the question: Y ~ N (10, 10^2): here. S9 pointed at the horizontal line he drew: and here is a square, 20 (pointed at the mean of X, as the	Self theory Square	5 61

width) and 20 (the standard deviation of X, as the height). Comparing to this, here is also a square, 10 (the mean of Y, as the width) and 10 (the standard deviation of Y, as the height). (S9 thought X ~ N (20, 20^2) and Y ~ N (10, 10^2) as Width, height and width, height of square		
I thought the height is 25, but you said it is 5^2, so I redrew it.	Variance and standard deviation	44
S9 thought standard deviation influence the height of curve	Self theory	5
S6 thought that 10 (mean of Y) is the x-coordinate and 10 (standard deviation of Y) is the y-coordinate	Self theory	5
L: why is the curve of Y narrower than the curve of X S6: because it is 10^2, and the first one is 20^2. L: why is 20^2 wider? S6: 20 is bigger, so the range is bigger. (this is contradict to his previous idea)	S.D bigger, curve wider	5
I thought it is 25^2, so I drew a very wide curve, but you remind us it is 5^2, not 25^2, so I drew a narrower curve.	Variance and standard deviation	44
S7: I feel that in the first graph, the x-ordinate is 20, and the height seems corresponding to 20. So in the second graph, the centre is 10, and I think that its height should be also 10. L: where do you get y-coordinate? S7 pointed at 10^2 of Y ~ N (10, 10^2) in the condition: here.	Focus on height (decided by S.D), not spread	5
L: the centres of second and third curves both locate at 10. S8: why 10?	Center decided by mean	10
It is shown that the content is 500c.c. on the package of a drink $H_0: \mu \leq 500 \quad H_1: \mu > 500$	=, > And <	46, 49
right-tailed testing $\alpha = 0.05, \quad Z_{0.025} = 1.96$ $Z_0 = \frac{498.2-500}{10/\sqrt{100}} = \frac{-1.8}{1} = -1.8 \ngtr -1.96$	Why 0.025 (influenced by confidence interval?) Why -1.96?	62 13

In Q6, $Z_0 = \frac{498.2-500}{10/\sqrt{100}} = \frac{-1.8}{1} = -1.8 \not> -1.96$ In Q7, $Z_0 = \frac{498.2-500}{10/\sqrt{100}} = -1.8 < -1.96$	Compare negative numbers	27 15
$H_0: \mu = 500$ $H_1: \mu < 500$ or $H_1: \mu \neq 500$ 2 tails testing  $H_0: \mu \geq 500$ $H_1: \mu < 500$ left-tailed testing  	Critical region the same; No test statistic	62 13 5
$H_0: \mu = 500$, $H_1: \mu < 500$	= And <	46
$H_0: \mu > 500$, $H_1: \mu < 500$ CR = {T < $t_{0.05}(99)$ = 1.671 ~ 1.658}	=, > And <; T distribution	46 5, 63
$Z_{0.05}$ = 1.645	One tail or two tails	49
The opposite of '=' is '≠'	=, ≠, > And <	64
S7: see 500 as the middle, 498.2 should be in the accept region. L: how do you know? S7: I guess. L: didn't you calculate? S7: how can I calculate without numbers?	Guess; Didn't see numbers	36 30
I solve them by myself… they are so easy that I don't have to memorise them.	Rely on self	36
I taught some classmates for several times, then may have impression of the answers.	Teach others	2
I memorise the answers because I don't know how and they are difficult.	Remember answers	13 36

There are plenty of *formulas*, and I don't remember what the *symbols* represent	Too many formulas; Symbols	65 48
its numbers are larger and I may enter a wrong number when using computer	Larger numbers	15
memorise the formulas, calculate by self.	Formula	13, 65, 36
I repeat solving them by myself, and I find that I memorise the processes and numbers naturally	Practice and remember	36 13
never learn…formulas in statistics are more complicated…using wrong formulas	Unfamiliarity Formula	30 65
statistics is very complex,…as difficult as mechanics,…I hate to see the symbols and numbers… I must have to remember many formulas, and I cannot solve the questions if I did not remember the formulas… cannot find the formula according to the conditions in questions, but may be able if using the book.	Hate symbols and numbers; Must remember formulas; Cannot find formulas by self	48 15 65 58
statistics is more flexible than calculus, because calculus just requires putting the numbers into the formula. It involves testing and software. It is not limited on paper work. I think it is more diversified.	Flexible; Diversed	66
many questions in calculus require you to calculate and get the answer, and many questions in statistics are words describing the condition and you have to find the correct formula by yourself	Difficulty; More than calculate	58 65
there are too many formulas.	Too many formulas	65
I am confused of the data of population and of sample…I often don't know the difference between population and sample…*To deal with statistics questions, you have to distinguish population and sample*…sometimes I regard large distribution as small distribution, or small distribution as large distribution. I don't know whether the 5th question is small or large distribution…I think it is *small [i.e. sample size]* because there are only *two departments*.	See 2 samples as sample size 2.	57 58
I sometimes get confused of estimation of population proportion p and estimation of difference	Too many formulas	65 13

of two population means $\mu_1 - \mu_2$. I cannot decide which it is according to the question. The formulas are really too many, so I cannot remember all of them and often get confused in problem solving. Statistics is also easy because you can get the result if you find and use the correct formulas. Sometimes I *remember some formulas* but I *forget what its condition is*.		
once I told teacher A to speed up. he may repeat the same type of question for 4 or 5 times	Too slow; Too many repeat	14
It depends on whether you listen to him or not.	Attitude	1
teacher A has different metaphor, such as he uses 'bad peer pretends to be apple' and 'seeing diamond as stone' to explain type I error and type II error. I think I will remember them forever. His descriptions are closer to life	Metaphor; Close to life	67 68 14
I like the atmosphere in his class.	Atmosphere	14
he usually said 'S5, you study so hard, I love you so much, please come to solve the question on the blackboard.' Also, he usually gets mix up and writes wrong answer on the blackboard.	Chat with students; Mistakes in class	14
he *explains in his words* first, rather than following the steps in the textbook. It makes me easier to understand…I cannot understand what the textbook is talking about, so I feel teacher A's method is good.	Easily understood words; Cannot understand; Textbook	14 47
he tried to help less capable students learn better.	Less capable	14
he usually gives us examples that are easy to understand, and it helps us learn faster.	Examples	14
I am uninterested in mathematics	Uninterested	1
He is a teacher so he knows what students are thinking. He uses some approaches closer to students to teach them.	Close to student	68
his explanation is easy to understand, he is close to students, and he uses simple approaches to teach us.		14
he always writes something on the blackboard first. It helps us understand statistics better because I cannot understand what the textbook is talking		14 47

about. After he explains in his words, I can realise what the textbook is about.		
he often makes wrong calculation and write wrong answer, this is what he has to improve.	Frequent mistakes	14
Practice question are aimed to increase our impression, to test how to use the formulas, and to prevent the fear of examinations. They also train our braveness because teacher A wants us to solve them on the blackboard.		69
they help us to know which formula should be used to solve each kind of question, and we know what formulas we should use in the examinations.	Formula and questions; familiarity	65 58
we have to think by ourselves	Real ‘practice’	69
He only lets us practice easy questions.	Easy	14
We are just taught a part of it and I don’t know what the later part is. Once my friend asked me what S.D. is, but I don’t know how to answer him.	Not learn all; Symbols	47 48
It is not only about ‘writing’, bust also ‘seeing’. Some people know what it is if they see the conditions. For example, I may forget sample S.D and population S.D soon.		58 57 48
I *cannot distinguish the conditions* when reading the questions.	Dinstinuish conditions	58
I only know how to use the formulas to find the answer. I don’t know how the formulas come, and even why there statistics is. I only know how to calculate and find the answer. I also don’t know how statistics can be used outside the classroom. It can be useful or useless in daily life.	Meaning(less) learning	66
statistics can be used for manage, test, measure, and…some charts, such as the ticket accumulation of election, and life of bulb. It is really frequently used in life, so I think statistics is useful. testing. it can be used in many circumstances, such as computer, design, and daily life. statistics can be used in daily life, such as total numbers, percentage. to see if someone exceeds the speed	Use of statistics	66

limit…gambling, probability		
I like it because I learn some knowledge to use; but I also don't like it because of its complexity. *Even if I was given the condition, I may still be unable to solve it…*you have to consider which formulas you should use according to the given conditions. The questions in a series are really hard. I like it because it may help me find a good job which requires many kinds of skill; I also don't like it because it may be totally useless in my entire life. I think I like it, because it can train our logic. Also, my memory of graph is not good, and I can train my ability of graph by drawing graphs by myself. I don't like it because it is very complicated, especially the conditions. I like it because it can be used in daily life; but I wish the data can be simplified	Feelings about statistics	66 58 65
S1 murmured: 1.96 L: do you remember that or do you find it in the textbook? S1: I remember that, that is enough. In Q9, it would be 1.96 if α is divided by 2.	Remember 1.96	13
$Z = \dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} = \dfrac{500-498.2}{\dfrac{10}{\sqrt{100}}} = \dfrac{0.18}{1} = 0.18$ It is true. [When I asked S2 to explain what $\overline{X}$ and μ mean, he wrote $\overline{X}$ above 500c.c. and μ above 498.2c.c; he used a correct formula and got a positive test statistic Z.] I thought *0.18 should compare to standard deviation…* smaller than standard deviation, so it is true.	Symbols; Self theory	48 5
In Q9, H_0 : μ = 500 H_1 : μ < 500 $Z_0 = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} = \dfrac{500-498.2}{10/\sqrt{100}} = 1.8$	Hypotheses; Symbols; Memory; Negative	46 48 13 27

[She reversely wrote $\overline{X}$ above 500c.c., and μ above 498.2c.c.; she used a correct formula and got a positive test statistic Z_0.] 90 $95 \rightarrow 0.05$ $99 \rightarrow 0.1$ $Z_{0.05} = \cancel{2.575}\ 1.96$ [Later she rewrote and got correct Z_0, she thought this is correct.] $Z_0 = \dfrac{498.2-500}{10/\sqrt{100}} = -1.8$ L: why do you have 2 sets of answers? S3: because I am *not sure about which* $\overline{X}$ *and* μ *are*. In Q10, $Z_{0.05} = 2.575$ [should be 1.645] S3: H_0: μ > 500, so it is *one tail testing*. α = 0.05, so the critical value is 2.575, but *it is negative because* Z_0 *is negative*.		
H_0: μ ≤ 100 H_1: μ > 100 $Z = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} = \dfrac{500-100}{\sqrt{498.2/10}}$ $= \dfrac{400}{\quad} ?\ Z_{0.05} \rightarrow 1.645$ L: what does μ mean? S4: is it *sample number*?	Hypotheses; Symbols; Memory; Self theory	46 48 13 5
In Q9, ~~n = 500~~ n=5 $\overline{X}$ = 498.2 σ = 10 H_0: μ = 498.2 H_1: μ ≠ 498.2 $Z_0 = \dfrac{498.2-500}{10/\sqrt{500}}$ (deleted) Z_0 =… = -1.8 < 1.96	Symbols; Hypotheses; Positive critical region	48 46 27 5 70

In Q710 $Z_0 = \ldots = -1.8 < 1.96$ L: why is $\mu = 498.2$ in H_0 ? S5: because it says 'gets the average 498.2c.c.'. L: why do you compare the negative result to 1.96? S5: if -1.8 is $\boldsymbol{\mu}_1$, and 1.96 is $\boldsymbol{\mu}_2$ **.** L: why is there $\boldsymbol{\mu}_2$ **?** S5 pointed at H_0 : $\mu = 498.2$: assume $\mu_1 = \mu_2$, H_1 : $\mu \neq 498.2$: assume $\mu_1 \neq \mu_2$. S5: pointed at $-1.8 < 1.96$: because $\mu_1 \neq \mu_2$, so reject H_0 . L: where does 1.96 come from? S5: significant level. L: significant level 0.05, so you get 1.96? S5: yes. L; how about Q10? S5: the same, I also compare μ_1 and μ_2 . L: why do you reject H_0 because -1.8 < 1.96? S5: $\mu_1 < \mu_2$ *(-1.8 < 1.96), so* $\mu_1 < \mu_2$ *(μ < 500).*		
S5compared -1.8 and +1.96. S1 compared -1.8 and -1.645 and he knew -1.8 < -1.645, so he got correct answer.	Negative; Critical value	27 70

6.21. Codes in the second and third student interviews.

NUMBER	CODES	INCLUDE	BELONG TO	LINK TO
1	Student issues	32, 37, 66, 68, 69	12	
2	Students' interaction			
3	**Key idea**	6, 17, 25, 27, 33, 35, 39, 42, 48, 49, 70		
4	**Difficulty**	17, 25, 27		12
5	Self theory	61	12	30
6	Area		3	
7	**Graph user**	60	36	
8	**Graph non-user**		36	
9	**Function of graph**		20	
10	Fundamental knowledge	44, 52, 53	12	
11	Discipline		5, 12	
12	**Cause of difficulty**	1, 5, 10, 13, 15, 19, 22, 29, 31, 43, 44, 46, 48, 49, 57, 58, 62, 64, 70, 71		4
13	**Memorise**	62	12, 36	
14	Class issues	47, 50, 67	12	
15	Numbers/ number line/ number sense		12	
16	Choice of table		36	
17	**Transfer/Principle**		3, 4	
18	Name		30	
19	Page arrangement		12	
20	Interpretation	9, 21, 23, 24		
21	Function of normalisation		20	
22	**Never consider**	27, 64	12	
23	Function/interpretation of confidence interval		20	
24	Definition of visualisation		20	
25	**> and <**		3, 4, 12	
26	**Follow**		36	
27	**Negative**		3, 4, 12, 22	
28	Format		12	
29	**Inner/outer of table**		12	
30	(Un)familiarity	18	12	5
31	**Mis-corresponding**		12	
32	**Solving orientation**		1, 36	
33	Not on the point		3	
34	Impressive			

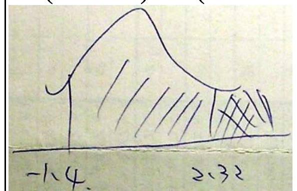
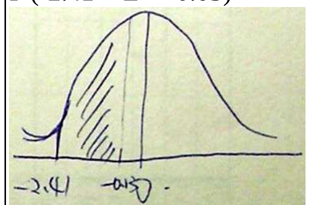
35	Between		3	
36	**Students' strategy**	7, 8, 16, 17, 26, 32, 42, 54, 60		
37	Experience		1	
38	Reverse thinking			
39	**Symmetry**		3	
40	Mis-interpreting	12	71	
41	Mis-calculation		71	
42	Outside-in & inside-out		3, 36	
43	(in)correctness of material		12	
44	Variance & S.D.		10, 12	
45	Mis-transcribing		71	
46	**Hypotheses**		12	
47	Content of material		14	
48	**Symbols**		3, 12	
49	Left/right tail & 2 tails		3	
50	Teacher		14	
51	Mis-transferring		71	
52	Mean		10	
53	Empirical rules		10	
54	**Compare and revise**		36	
55	Mis-targeting		71	
56	Mis-reading		71	
57	Statistic language		12	
58	**Conditions**		12	
59	First digits		12, 40	
60	Unmatched graph		7, 36	
61	Square		5	
62	**From α to z score**		12, 13	
63	T distribution			
64	$\neq$		12, 22	
65	**Too many formulas**		12	
66	Feeling of statistics		1	
67	Metaphor		14	
68	Close to life		1	
69	Practice questions		1	
70	Critical value		3, 12	
71	Mistakes of using table	40, 41, 45, 51, 55, 56	12	

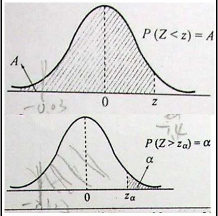
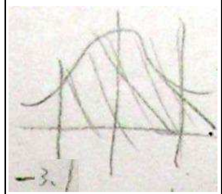
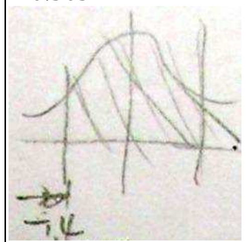
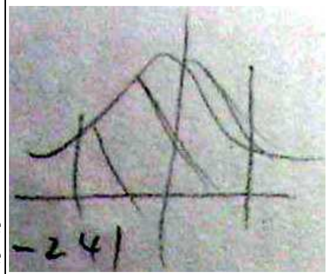
6.22. Categories in the second and third student interviews.

CATEGORIES	INCLUDING CODES	SUCH AS
Key Ideas (Must-know & Factor)	Area	
	Transfer & Principle	
	> and <, = and ≠	
	Negative	
	Between	
	Symmetry	
	(Not) on the point or table	
	Inner/outer of table / Outside-in & Inside-out	
	From $(1 – \alpha)\%$ to z score / Confidence intervals	
	Hypotheses & Left / right tail & 2 tails	
	From α to z score / Critical values	
	Symbols	
	Numbers / Number line	
Difficulty & Cause of Difficulty	Self theory	
	Transfer / Principle	
	> and <	
	Negative	
	Not on the point	
	Format	$P(Z > a) = 0.67$, $P (Z > 0.67) = b$
	Inner/outer of table	
	Student issues	Attitude; Experience;
	Fundamental knowledge	Mean; Variance & S.D; Empirical rules;
	Numbers / Number line	
	Mis-corresponding (homography)	
	Mistakes (of using table)	Mis-reading; Mis-transcribing; Mis-transferring; Mis-targeting; Mis-interpreting; Mis-calculation

	Memory	
	Unfamiliarity	
	Never consider	≠; Negative
	From conditions to formulas	Statistic language; Too many formulas
	Hypotheses & Left/right tail & 2 tails	
	From α to z score	
	Symbols	
	First digits (.0505 & .0516 for .05, not .4995)	
	Class issues	Teacher's mistake
	Material (textbook)	Page arrangement; Incorrectness; Content
Students' Strategies	Graph users	(un)matched graphs
	Graph un-users	
	Choice of tables	
	Transfer/Principles	
	Follow	
	Memorise	
	Solving orientation	
	Outside-in & Inside-out	
	Compare and revise	
	Square-like	
Students' Interpretation	Definition of visualisation	
	Function of graph	
	Function of normalisation	
	Function/definition of confidence interval	
	Feeling, understanding & use of statistics	
Others **(Rarely Mentioned)**	Student interaction	
	Impressive of the process of realizing	
	T distribution	
	Class issues	Teacher's metaphor; Close to life; Atmosphere; Teaching style; Practice questions

7.1. Correct answers and the students' solutions (with explanations) in Q1.1-Q1.6.

Question No.	Q1.1	Q1.2	Q1.3	Q1.4	Q1.5	Q1.6
Question	P (Z < -0.03)	P (Z < 1.22)	P (Z > -3.1)	P (Z > 2.05)	P (-1.4 < Z < 2.32)	P (-2.41 < Z < -0.03)
Correct Answer	0.488	0.8888	0.999	0.0202	0.9898 - 0.0808 = 0.909 (or 1–0.0808–0.0102)	0.488 – 0.008 = 0.48 [or (1 - 0.008) - (1 - 0.488)]
S1 (prefers Table A, but was given Table B) XOXOXXX	P (Z < -0.03) = 0.5120	P (Z < 1.22) = 0.8888	P (Z > *3.1*) = 0.0010	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) ↓ = 0.5667 0.898 = 0.3231 [0.5667 is from 1 – 0.4443 which corresponds to.14, and 0.898 is from 1 – 0.0102 which corresponds to 2.32]	P (-2.41 < Z < -0.03 = 0.5040
S2 (prefers Table B, and was given B) OXXXXXX	-0.03; 0.4880	< 1.22; 0.1112	> - 3.1; *0.010*	> 2.05; 0.202	-1.4 < Z < 2.32; *0.0808* < Z < 0.0102	-2.41 < Z < -0.03; 0.0080 < Z < 0.4880
S3 (Table B) S3 - S10 were given both Table A and Table B. The table in ( ) means their preference OOOOXXO	P (Z < -0.03) = 0.4880	P (Z < 1.22) = 0.8888	P (Z > -3.1) = 1 – P (Z > 3.1) = 1 – 0.0010 = 0.999	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) =P (Z > -1.4) – P (Z > 2.32)  = 1 – P (Z > 1.4) – P (Z > 2.32) = 1 - 0.0808 - *0.0104* = *0.088*	P (-2.41 < Z < -0.03)  = P (Z > -2.41) – P (Z > *-0.13*) = 1 – P (Z > 2.41) – [1 – P (Z > 0.13)] = Z > 0.13 – P (Z > 2.41) = 0.4483 - 0.0080 = 0.4403
S4 (Table A) XXXXXXX	P (Z < -0.03) = 0.5120	P (Z < 1.22)	P (Z > -3.1) = 0.0010	P (Z > 2.05) = 0.9798	P (-1.4 < Z < 2.32)	P (-2.41 < Z < -0.03)

S5 (Table B) OOXXXXX	P (Z < -0.03) = ~~*0.4920*~~ = 0.4880	P (Z < 1.22) = 0.8888	P (Z > *3.1*) = *0.0010*	P (Z > 2.05) = *0.0207*	P (-1.4 < Z < 2.32) = P (Z > -1.4) – P (Z > 2.32) = 1 – P (Z > 1.4) – P (Z > 2.32) = 1 - 0.0808 - *0.0104*	P (Z > -2.41) – P (Z > -0.03) = 1– P (Z > 2.41) – 1 – P (Z > 0.03) = 1 – *0.0082* – 1 – 0.4880
S6 (Table B) OOOOOOO	P(Z < -0.03) = P (Z > 0.03) = 0.488  [S6 directly marked -0.03 on the graphs in the textbook]	P (Z < 1.22) = 0.8888	P (Z > -3.1) = 1 – P (Z > 3.1) = 1-0.001 = 0.999 	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) = P (Z > -1.4) – P (Z > 2.32) = 1 – P (Z > 1.4) – P (Z > 2.32) = 1 - 0.0808 - 0.0102 = 0.909  [This graph is the same to the graph in Q1.3, S6 just replace -3.1 by -1.4.]	P (-2.41 < Z < -0.03) = P (Z > -2.41) – P (Z > -0.03) = 1 – P (Z > 2.41) – 1 + P (Z > 0.03) = 1 – 0.008 – 1 + 0.488 = 0.48 
S7 (Table B, but used Table A to check P(Z<-z) OOOOOXO	P (Z < -0.03) = P (Z > 0.03) = 0.488	P (Z < 1.22) = 1 – P (Z > 1.22) = 1 - 0.1112 = 0.8888	P (Z > -3.1) = P (Z > -3.1) = 1- P (Z > 3.1) = 1 - 0.0010 = 0.999	P (Z > 2.05) = P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) = P (Z < 2.32) - *P (Z > -1.4)* = 1 – P (Z > 2.32) – *P (Z > 1.4)* = 1 - 0.0102 - 0.0808 = 0.909	P (-2.41 < Z < -0.03) = P(Z < -0.03) – P(Z > -2.41) ~~= *1 – (P > 0.03)* – P(Z > 2.4)~~ ~~= 1 – 0.488 – 0.0082~~ ~~= 0.5038~~ => P (Z > 0.03) – P (Z > 2.4) = 0.488 - 0.0082 = 0.4798
S8 (Table B) OXOOXXX	P (Z < -0.03) = P (Z > 0.03) = 0.4880	P (Z < 1.22) = P (Z > 1.22) = 0.1112	P (Z > -3.1) = P (Z > 3.1) = 0.9990	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) =	P (-2.41 < Z < -0.03) =

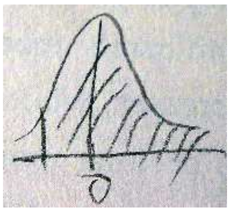
S9 (Table B) XXXOOOX	P (Z < -0.03) =P($\frac{x-\mu}{\sigma_x}$ < -0.03) = P (x < -0.03) = ~~1 –~~ P (x > 0.03) = ~~1 -~~ 0.4880 ~~= 0.5120~~ [The initial answer was 0.5120. S9 deleted “1-” in last third and second columns and left 0.4880 as the answer.]	P (Z < 1.22) = P($\frac{x-\mu}{\sigma_x}$ < 1.22) = P (x > 1.22) = 0.1112	P (Z > -3.1) = P($\frac{x-\mu}{\sigma_x}$ > -3.1) = P (x > -3.1) = 1 – P (x > 3.1) = 1 - 0.0010 = 0.9989	P (Z > 2.05) = P (x-μ/σ > 2.05) = P (x > 2.05) = 0.0202	P (-1.4 < Z < 2.32) = P (x > -1.4) – P (x > 2.32) = 1 – P (x > 1.4) – P (x > 2.32) = 1 - 0.0808 - 0.0102 = 0.9192 - 0.0102 = 0.909	P (-2.41 < Z < -0.03) = P (x > -2.41) – P (x > -0.03) = 1 – P (x > 2.41) – 1 – P(x > 0.03) = (1 - 0.008) - (1 - 0.488) = 0.992 - 0.512 = 0.48
S10 (Table B & negative part of Table A) OOOOOXX	P (Z < -0.03) = 0.488	P (Z < 1.22) = 1 – P (Z > 1.22) = 1 - 0.1112 = 0.8888	P (Z > -3.1) = 1 – P (Z > 3.1) = 1 - 0.001 = 0.999	P (Z > 2.05) = 0.0202	P (-1.4 < Z < 2.32) = 1 - P (Z > 2.32) - *(Z < -1.4)* = 1 - 0.0102 - 0.0808 = 0.909	P(-2.41 < Z < -0.03) = P (Z > 0.03) - *(Z <-2.41)* = 0.488 - *0.0207* = 0.4673 [0.0207 = P (Z < -2.04)]

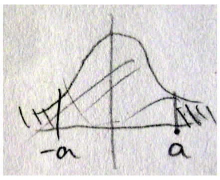
Explanation of the table:

Italics means wrong place. Solutions deleted by students are displayed and shadowed.

In the first column, O means that the final answer is correct and X means the final answer is wrong, For example, S1's XOXOXXX means his answer of Q2 & Q4 are correct, and others are wrong (mistakes in the solving process are neglected in this table).

7.2. Correct answers and the students' solutions (with explanations) in Q1.a-Q1.d.

Question No.	Q1.a	Q1.b	Q1.c	Q1.d
Question	P (Z > a) = 0.67 Find a	P (Z > 0.67) = b Find b	P (Z < c) = 0.67 Find c	P (Z < 0.67) = d Find d
Correct answer	**a = -0.44** **Solving by Table A:** **P(Z < 0.44) = 0.67** **P(Z > -0.44) = 0.67** **Solving by Table A:** **P (Z < a) = 0.33** **Solving by Table B:** **P(Z > -a)= 1–P(Z>a)** **=0.33** **-a = 0.44** **a = -0.44**	**b = 0.2514**	**c = 0.44**	**d = 0.7486**
S1 (Table A & *B*)	a = -0.44	b = 0.2514	c = 0.44	d = 0.7486
S2 (Table A)	0.44			
S2 supplemented (Table A)		~~0.2206~~ 0.2514	1 - 0.67 = 0.33 c = 0.6293	
S3 (Table A & B)	~~1 - P(Z > a) = 1 - 0.67~~ P(Z<-a)= 1-0.67 = 0.33 -a = -0.44 a = 0.44  [S3 retried.] P(Z > a) = 0.67 1 - P(Z > a) = 0.67 P(Z > a) = 0.33 a = 0.44	b = 0.2514	c = 0.44	d = 0.7486
S4 (Table A)	a = 0.2514			
S5 (Table A & B)	P (Z > a) = 0.67 a = ~~0.43~~ 0.2514	P (Z > 0.67) = ~~0.7486~~ 0.2514	P (Z < c) = 0.44 or 0.45	P (Z < 0.67) = 0.7486
S6 (Table A &	P(Z > a) = 0.67	P (Z > 0.67) = b	P (Z < c) = 0.67	P (Z < 0.67) = d

B)	=> 1 – P(Z<a) = 0.67 => P (Z<a) = 0.67 a = -0.44 	=> b = 0.2514	c = 0.44	= 0.7486 = d
S7 (Table A & B)	P(Z > a) = 0.67 **a** => ~~0.44~~ ~~1.49 < a ≤ 1.50~~ 1.49 > a ≥ 1.50	P(Z>0.67) = 0.2514	P (Z < c) = 0.67 c = 0.44	P(Z < 0.67) = d d = 0.7486
S8 (Table A & B)	1 – 0.67 = 0.33 a = ~~0.04~~ 0.404	0.2514	P(Z < 0.33) = ~~0.6255~~ 0.6293	0.7486
S9 (Table A & B)		0.2514		-0.2514
S10(Table A& B)	-0.44	0.2514	0.44	0.7486
Correct: Wrong	3: 7	8:2	5:5	7:3

7.3. Counting of Table 7.2, Table 7.3 and Table 8.3.

Table 7.2: Q1.1-Q1.6

Difficulties & Causes \ Stage			Question	Information	Interpretation	Action
Symptoms	Mis-reading		-	-	2 (Q1.3: S1,S5)	-
	Mis-transcribing		-	-	-	6 (Q1.5: S10; Q1.6: S3, S3, S7, S7, S10)
	Mis-fitting		-	-	5 (Q1.1: S4 Q1.2: S2 Q1.3:S2, S4; Q1.4: S4)	0
	Mis-transferring		-	-	-	8 (Q1.1: S9; Q1.2: S8, S9; Q1.3: S8; Q1.5: S7, S7; Q1.6: S7, S7)
	Mis-targeting		-	-	-	5 (Q1.1: S5; Q1.4: S5; Q1.5: S3, S5; Q1.6: S5)
	Mis-interpreting		-	-	-	4 (Q1.3: S2; Q1.4: S2; Q1.5: S1; Q1.6: S10)
	Mis-calculation		-	-	-	5 (Q1.3: S9; Q1.5: S1, S1, S3; Q1.6:S5)
Causes	Reading the question in symbolic form	Relation sign	-	-	4 (Q1.1: S2, S8; Q1.2: S2, S8)	3 (Q1.1: S9; Q1.2: S9; Q1.6: S7)
		Minus sign	-	-	8 (Q1.1: S2, S4, S8; Q1.3: S1, S2, S4, S5, S8)	4 (Q1.1: S9; Q1.5: S7; Q1.6: S7, S7)
		When Z is between two values	-	-	-	11 (Q1.5: S1, S2, S4*, S7, S8*; Q1.6: S1, S2, S4*, S5, S7, S8*)
		Question format	-	-	12 (Relation sign + Minus sign)	7 (Relation sign + Negative number)
	Using the table(s)		-	-	-	9 (Mis-targeting + Mis-interprting)

	When the answer was not in the table	-	-	9 (Q1.1: S2*, S8* Q1.2: S2, S4*, S8; Q1.3: S2,S4,S8; Q1.4: S4)	9 (Q1.1: S1, S9; Q1.2: S9; Q1.5: S1, S2, S7; Q1.6: S1, S2, S7)

* *Numbers in the table means how many times such situation happened; repeated symptoms in the same question were counted multiple times; '-'means inappropriate.*

** *I categorise S8 in Q1.3 [P(Z > -3.1) = P(Z > 3.1) = 0.001] in both 'mis-transferring' and 'minus sign', but put S2 in Q1.3 [P(Z > -3.1) = 0.01] only in 'minus sign'. S9 is counted in 'action' rather than 'interpretation' because he explained that he mistakenly memorised the transferring principles.*

*** *S2 and S8 had correct answer in Q1.1, but they are categorised to 'directly using Table B in the stage of interpretation' by comparing to their solutions in other sub-questions.*

*****S4 did not answer Q1.2, Q1.5 and Q1.6. S8 did not answer Q1.5 and Q1.6.*

Table 7.3: Q1.a-Q1.d

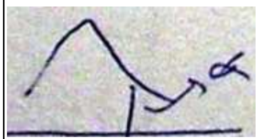
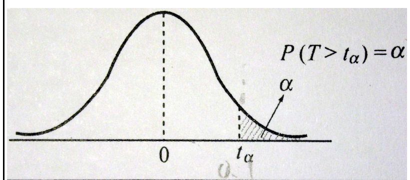
Difficulties & Causes \ Stage			Question	Information	Interpretation	Action
Symptoms	Mis-reading		-	-	0	-
	Mis-transcribing		-	-	-	0
	Mis-fitting		-	-	4 (Q1.a: S2, S4, S5; Q1.b: S5)	0
	Mis-transferring		-	-	-	4 (Q1.a: S3, S3; Q1.c: S8 Q1.d: S9)
	Mis-targeting		-	-	-	4 (Q1.a: S5; Q1.b: S2; Q1.c: S5, S8)
	Mis-interpreting		-	-	-	4 (Q1.a: S5, S7, S8; Q1.c: S2)
	Mis-calculation		-	-	-	0
Causes	Reading the question in symbolic form	Relation sign	-	-	6 (Q1.a: S2, S5, S7, S9; Q1.b: S5; Q1.c: S2)	3 (Q1.a: S3, S8; Q1.d: S9)
		Minus sign	-	-	-	-
		When Z is between two values	-	-	-	-
		Question format	-	-	17 (Q1.a: S2, S4, S5, S5, S7, S9; Q1.b: S2, S4, S5; Q1.c: S2, S2, S4, S8, S9; Q1.d: S2,S4,S9)	5 (Q1.a: S3, S8; Q1.c: S2, S8; Q1.d: S9)
	Using the table(s)		-	-	-	8 (mis-targeting + mis-interpreting)
	When the answer was not in the table (Negative answer)		-	-	3 (Q1.a: S2,S7,S8)	1 (Q1.a: S3)

Table 8.3: Q3.1-Q3.2

Difficulties & Causes \ Stage		Question	Information	Interpretation	Action
Symptoms	Mis-reading	-	-	0	-
	Mis-transcribing	-	-	-	0
	Mis-transferring	-	-	-	5 (Q3.2: S4,S5,S6,S8,S10)
	Mis-targeting	-	-	-	0
	Mis-interpreting	-	-	-	3 (Q3.1: S1, S1; Q3.2: S4,)
	Mis-calculation	-	-	-	0
Causes	Symbolic language (Question format)	-	-	10 (Q3.1: S7; Q3.2: S2*, S3*, S4, S5, S6, S7, S8, S9*, S10)	7 (Q3.1: S7; Q3.2: S4,S5, S6, S7, S8, S10)
	Using the table (The degree of freedom)	-	-	-	2 (Q3.1: S7; Q3.2: S7)
	Not in the table (Negative answer)	-	-	8 (Q3.2: S2*, S3*, S4, S5, S6, S8, S9*, S10)	5 (Q3.2: S4, S5, S6, S8, S10)

**S2, S3 and S9 did not answer Q3.2*

8.1. Correct answers and the students' solutions (with explanations) in Q3.1-Q3.2.

Question No.	Q3.1	Q3.2
Question	$t_{0.005}(9)$	$t_{0.99}(23)$
Correct Answer	**3.25**	**– 2.5**
S1	$t_{0.005}(9)$ = ~~6.594~~ [which is $t_{0.00005}(9)$] ~~4.781~~ [which is $t_{0.0005}(9)$] 3.250	$t_{0.99}(23)$ = – 2.5000
S2	$t_{0.005}(9)$ = 3.250	$t_{0.99}(23)$ cannot find
S3	$t_{0.005}(9)$ = 3.250	$t_{0.99}(23)$ = 
S4	$t_{0.005}(9)$ = 3.250	$t_{0.99}(23)$ = 1.319 [which is $t_{0.1}(23)$]
S5	$t_{0.005}(9)$ = 3.250	$t_{0.99}(23)$ = 1 – 0.01 = 1 – 2.500
S6	$t_{0.005}(9)$ = 3.25	$t_{0.99}(23)$ = 31.821 – 2.5 [which is $t_{0.01}(1) - t_{0.01}(23)$] = 29.321 [S6: I guess.] 
S7	$t_{0.005}(9)$ = 3.355 [S7 thought v = 9 – 1 = 8.]	$t_{0.99}(23)$ = – 2.508 [S7 thought v = 23 – 1 = 22.]
S8	$t_{0.005}(9)$ = 3.250	$t_{0.99}(23)$ = –1.500 1 – 0.99 = 0.01 1 – 2.500 = –1.500
S9	$t_{0.005}(9)$	$t_{0.99}(23)$
S10	$t_{0.005}(9)$ = 3.25	$t_{0.99}(23)$ = 31.821 – 2.5 – 2.5 = 26.821 (S10 guessed)

10.1. Approaches of hypothesis testing.

Critical –value (classical test)

In the approach of critical-value, the critical-value is decided by the significant level α. Wonnacott and Wonnacott (1990) notice that individuals often confuse statistical significance with ordinary significance, and they stress that

> statistically significant is a technical phrase that simply means enough data has been collected to establish that a difference does exist (1990, p.291).

Critical region is also decided by whether it is a right-tailed, left-tailed or two-tailed testing. Therefore, there will be three kinds of critical value,

critical value = $\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$ (right-tailed);

$\mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ (left-tailed); or

$\mu_0 \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (two-tailed).

The null hypothesis H_0 will be rejected only when the observed sample mean $\overline{X}$ falls in the critical region. That is, when $\overline{X}$ is bigger than $\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$ in right-tailed testing, smaller than $\mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ in left-tailed testing, or bigger than $\mu_0 + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ or smaller than $\mu_0 - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ in two-tailed testing. Such an approach is also called classical test.

Test statistic

This method is similar to critical-value approach. However, no matter if the hypothesis is left-tailed, right-tailed or two-tailed, the test statistic of mean remains $Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$.

The null hypothesis H_0 will be rejected only if the test statistic falls in the rejection region. The rejection region is decided by the type of hypothesis. To be specific, the rejection region in left-tailed testing is $\{Z < -Z_\alpha\}$, the rejection region in right-tailed testing is $\{Z > Z_\alpha\}$, and the rejection region in two-tailed testing is $\{Z < -Z_{\frac{\alpha}{2}}$ or $Z > Z_{\frac{\alpha}{2}}\}$.

Confidence interval

In the previous category, confidence interval was introduced as the result of interval estimation. In fact,

> it may be regarded as the set of acceptable hypotheses… Once a confidence interval has been calculated, it can be used immediately to test any hypothesis… by examining whether or not it (the null hypothesis, such as μ_0) falls within the confidence interval (Wonnacott and Wonnacott, 1990, pp.289, 291, 293).

Students' experience of confidence interval may help them apply this method to do two-tailed testing. However, their experience of symmetric confidence interval may not help when they do one-tailed testing.

There is an equivalence of the two-sided critical region approach and two-sided confidence interval approach. In the critical region approach, the critical regions are outside the acceptance region which is the middle chunk centred around the assumed population mean μ_0, and H_0 will be rejected if the observed sample mean $\overline{X}$ falls in the critical region. In the confidence interval approach, the confidence interval is centred around the observed sample mean $\overline{X}$, and H_0 will be rejected if the assumed population mean μ_0 falls outside the confidence interval. The length of the acceptance region is equal to the length of the confidence interval. Taking significant level 5% for instance, the critical regions are edged at $\mu_0 \pm 1.96\frac{\sigma}{\sqrt{n}}$, and the 95% confidence interval of mean is $\overline{X} \pm 1.96\frac{\sigma}{\sqrt{n}}$. Therefore, when $\overline{X}$ falls outside the acceptance region $(\mu_0 - 1.96\frac{\sigma}{\sqrt{n}}, \mu_0 - 1.96\frac{\sigma}{\sqrt{n}})$, μ_0 also falls outside the confidence interval $(\overline{X} - 1.96\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma}{\sqrt{n}})$. That is, in both approaches, we check "whether the magnitude of the difference $\left|\overline{X} - \mu_0\right|$ exceeds 1.96 standard errors". The difference between the two methods is that the critical region approach uses the assumed population mean μ_0 as its reference point while the confidence interval uses the sample mean $\overline{X}$ as its reference point (Wonnacott and Wonnacott, 1990, pp.316-317).

Moreover, one-sided confidence interval is also equivalent to the one-sided critical region approach, but one-sided confidence interval was not taught in the participants' course, so it was not used by the participants.

p-value

The last approach is to concentrate on the only one null hypothesis H_0, and to calculate the p-value as "how little it [i.e. H_0] is supported by the data" (Wonnacott and Wonnacott, 1990, p.293). P-value is a good way to "summarise what the data says about the credibility of H_0" (1990, p.295). When the alternative hypothesis H_1 is on the right-hand side, the p-value can be defined as the probability of that

> the sample value would be [at least] as large as the value actually observed, if H_0 is true (Wonnacott and Wonnacott, 1990, p. 294).

On the other hand, when H_1 is on the left side, the p-value is the probability of that

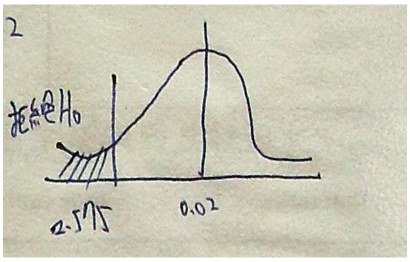
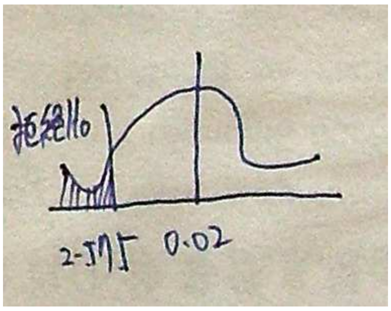
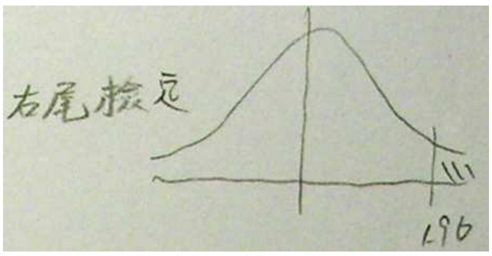
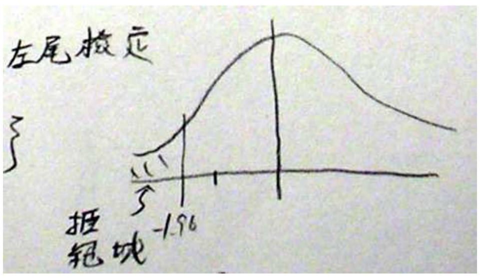
> the sample value would be [at most] as small as the value actually observed, if H_0 is true (Wonnacott and Wonnacott, 1990, p. 294).

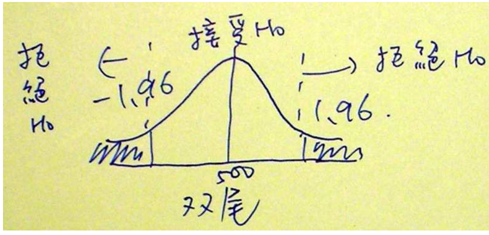
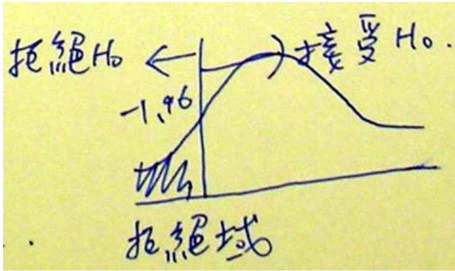
Therefore, such a method is a good way to "summarise what the data says about the credibility of H_0 … If this credibility falls below α, then H_0 is rejected" (Wonnacott and Wonnacott, 1990, pp. 293-294, 301). In this method, one-sided p-value is simply called p-value for brevity and convenience, and two-sided p-value is not abbreviated (Wonnacott and Wonnacott, 1990, p.314).

There are more and more applied statisticians choosing p-value method, because the other methods involve "setting α arbitrarily" and p-value method "avoids the arbitrary 95% or 5% level" since the p-value is calculated from the observed $\overline{X}$ (Wonnacott and Wonnacott, 1990, pp. 300, 302, 318). Beside, in the statistics courses adopting computer software, p-value becomes the critical method. Taking hypothesis testing in SPSS for example, the users can easily get the p-value by entering data and clicking options without manual calculation, and then compare it to the significant level α such as 5%. The null hypothesis will be rejected if the p-value is below the significant level.

10.2. Correct answers and S1-S10's solutions in Q6 and Q7 in the second interview.

	Q6 (two-tailed)	Q7 (left-tailed)
Correct answer	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ = 500** H_1 **: μ ≠ 500 (two-tailed testing)** $Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$ **-** $Z_{\frac{\alpha}{2}}$ **= -1.96** **CR = {Z > 1.96 or Z < -1.96}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 > -1.96** **Accept** H_0**, the content is 500c.c.**	**n = 100** $\overline{X}$ **= 498.2** **s = 10** **α = 0.05** H_0 **: μ ≥ 500** H_1 **: μ < 500 (left-tailed testing)** **-Zα = -Z0.05 = -1.645** **CR = {Z < -1.645}** $Z_0 = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $= \dfrac{498.2-500}{\frac{10}{\sqrt{100}}}$ **= -1.8 / 1** **= -1.8 < -1.645** **Reject** H_0**, the drink seller is not telling the truth.**
S1	H_0 = 500 H_1 < 490 500 × 0.05 = 25 > 10 So reject	H_0 < 500 H_1 > 500 Standard = 10c.c. Average 498.2c.c. H_1 is confirmed. True.
S2	standard deviation 10; examine and get 498.2c.c. ~~not exceeding standard deviation 10c.c., so they all are~~ then in the range of standard deviation, So it is true.	μ_0 = 500, n = 100, $\overline{X}$ = 498.2, s = 10, α = 0.05 H_0 : μ > 500, H_1 : μ< 500
S3	H_0 **:** μ ≥ 500 n = 100 **α = 0.05** H_1 **:** μ < 500 σ = 10 P ($\overline{X}$ > 498.2 / μ ≥ 500) = P ($\overline{X}$ - μ > 498.2 - 500)	

	$= P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{-0.8}{10/\sqrt{100}}\right)$ = P (Z > -0.8) = P (Z < 0.8) = 0.7881 Z= 1.96	
S4		
S5	H_0 : μ = 500 H_1 : μ < 500 $Z = \frac{498.2-500}{\frac{10}{100}}$ $= \frac{-0.02\times 10}{100}$ $= \frac{0.2}{100}$ $= \frac{20}{1000}$ $= 0.02$ CR = {Z > 2.575} 0.02 < 2.575 reject H_0 	**H_0 : μ ≥ 500** **H_1 : μ < 500** $Z = \frac{498.2-500}{\frac{10}{100}}$ = 0.02 CR = {Z > 2.575} 0.02 < 2.575 Reject H_0 
S6	H_0 : μ ≤ 500 H_1 : μ > 500 right-tailed testing α = 0.05 $Z_{0.025} = 1.96$ CR = {Z > $Z_{0.025}$ } 	**H_0 : μ ≥ 500 H_1 : μ < 500** left-tailed testing CR = {Z < - $Z_{0.025}$ }  $Z_0 = \frac{498.2-500}{10/\sqrt{100}}$ = -1.8 < -1.96

	$Z_0 = \dfrac{498.2-500}{10/\sqrt{100}}$ $= \dfrac{-1.8}{1}$ $= -1.8 \ngtr -1.96$	Accept H_0 [S6 did not write down the decision until I asked him to.]
S7	$H_0: \mu = 500$ $H_1: \mu < 500$ or $H_1: \mu \neq 500$ 2 tails testing  (because H_0 is true, the result is to 'not reject H_0') judgment correct, it is true accept H_0	$H_0: \mu \geq 500$ $H_1: \mu < 500$ left-tailed testing  H_0 is fake, the result is to reject H_0 judgment correct, it is true
S8	$\sigma = 10, \alpha = 0.05, Z_{0.05} = 1.645,$ $n = 100, \overline{X} = 498.2$ $H_0: \mu = 500$ $H_1: \mu < 500$ CR = {Z > $Z_{0.05}$ = 1.645}	$H_0: \mu \geq 500$ $H_1: \mu < 500$
S9	$\mu_0 = 500cc, n = 100, \overline{x} = 498.2cc,$ $s = 10cc, \alpha = 0.05$ $H_0: \mu \geq 500$ $H_1: \mu < 500$ CR = {T < $t_{0.05}(99)$ = 1.671 ~ 1.658}	$\mu_0 > 500cc, n = 100, \overline{x} = 498.2cc,$ $s = 10cc, \alpha = 0.05$ $H_0: \mu \geq 500$ $H_1: \mu < 500$ CR = {T < $t_{0.05}(99)$ = 1.671 ~ 1.658}
S10	$H_0: \mu > 500$ $H_1: \mu < 500$ $Z_{0.05}$ = 1.645	