*Jurnal Pendidikan Malaysia 40(2)(2015): 193-196*

# An Overview of Students' Learning Problems in Hypothesis Testing
(Satu Gambaran Masalah Pembelajaran Pelajar dalam Ujian Hipotesis)

Saras Krishnan* & Noraini Idris

ABSTRACT

*Hypothesis testing is an important tool of statistical inference and is taught in most introductory Statistics Courses. However, students' continuing struggle in making deep and connected understanding of the statistical concepts in the learning of hypothesis testing and its' use is evidential from past studies. This article presents an overview of the problems faced by students in the learning of hypothesis testing. Understanding students' learning of statistics in the introductory statistics courses are needed especially because introductory statistics courses have been the focus of statistics reform and because there has been great increase in the number of students in these courses in the recent years.*

*Keywords: Conceptual understanding; hypothesis testing procedure; misconceptions*

ABSTRAK

*Ujian hipotesis merupakan satu alat penting dalam inferens statistik dan diajar dalam kebanyakan kursus Pengenalan Statistik. Walau bagaimanapun, kajian sebelum ini membuktikan bahawa para pelajar masih menghadapi masalah dalam mencapai pemahaman yang dalam dan berhubung kait akan konsep-konsep statistik dalam pembelajaran ujian hipotesis dan penggunaannya. Artikel ini memberikan satu gambaran tentang masalah yang dihadapi oleh para pelajar dalam pembelajaran ujian hipotesis. Pemahaman tentang masalah yang dihadapi oleh pelajar-pelajar dalam pembelajaran statistik dalam kursus Pengenalan Statistik adalah diperlukan. Kursus Pengenalan Statistik merupakan fokus reformasi statistik dan kerana bilangan para pelajar dalam kursus ini telah meningkat sejak kebelakangan ini.*

*Kata kunci: Pemahaman konsep; prosedur ujian hipotesis; konsepsi salah*

## Introduction

According to Sotos et al. (2007), students' various misconceptions in the learning of inferential statistics are one of the important concerns of research in inferential statistics. Moreover, inferential statistics is an important topic in statistics courses for a majority of disciplines (Sotos et al. 2007). This has led to statistics educators and researchers studying diverse aspects of students' learning of inferential statistics. Among the different topics of inferential statistics, hypothesis testing is taught in almost every course encompassing different levels and students from different majors. This is because hypothesis testing is frequently used in a statistical analysis concerning population.

Hypothesis testing essentially involves making a claim about the population and testing the claim by analysis of sample data through a set of procedure. Hypothesis tests can include one-tailed test involving the population mean or the population proportion and the two-tailed test involving the population mean or the population proportion. The extent to which students learn hypothesis testing depends on their course syllabus and program of studies. Besides knowing the procedure of conducting hypothesis testing, students must have sound knowledge of the underlying concepts such as the sampling distribution and the related central limit theorem. Students must also have the fundamental statistical knowledge like the ability to differentiate sample and population and to differentiate population mean and population proportion.

## Research in Hypothesis Test

Researchers construed that hypothesis test is the main tool in inferential statistics and it is one of the most difficult topic to teach (Garfield & Ben-Zvi 2008; Link 2002; Sotos et al. 2007) especially in the introductory statistics courses (Link 2002; Zieffler et al. 2008). Also, researchers agree that even when students are able to perform the multi-step procedure of a hypothesis test, they cannot perceived the rationale of executing these steps or applying them in novel contexts. In his study, Link (2002) observed hypothesis testing as a six-step procedure which comprised of stating the null hypothesis and alternate hypothesis, calculating the critical value, forming the probability statement, calculating the observed value, stating the p-value and arriving at a decision pertaining to the problem in concern. This six-step procedure, Link (2002) reasoned is indicative of students' understanding whereby incorrect responses for individual items implied that students do not have

complete understanding of hypothesis test. However, the study did not explore the students' different levels of understanding.

Evangelista and Hemenway (2002) used a cooperative learning technique known as Jigsaw in order to help students make their own generalizations of hypothesis testing because they believe that group activities can enhance students learning of many statistical concepts and as an alternative method in guiding students to make connections in testing various parameters. On the other hand, Lipson, Kokonis and Francis (2003) used a web-based computer simulation to help students' form conceptual understanding of the role of sampling distribution in hypothesis testing. In an ensuing study, they further investigated the design features of the interactive simulation that enable students to gain maximum benefit from it (Lipson et al. 2006). They feel that instructors must work closely with students to enhance the students' learning experience in an interactive technology environment.

Zieffler et al. (2008) analyzed the meaning of informal inferential reasoning to provide a framework for studies on students' reasoning about statistical inference. They were concerned with components of the framework that supports its purpose and also the suitable type of tasks. Following this, Weinberg, Wiesner and Pfaff (2010) used this framework for informal inference in designing a hands-on activity to help students understand concepts of confidence interval and hypothesis test. This was achieved through informal reasoning, connecting this reasoning to formal ideas, and encouraging students to conjecture, explain and reflect their reasoning and ideas. It was also mentioned that instructors can use this activity to teach other important statistical concepts incorporated in it such as sampling variation and sampling distribution.

Smith (2008) argued in her doctorate dissertation that in analyzing students' understanding of hypothesis test, it is important to address the students' abilities to coordinate various ideas and ways of reasoning as well as their abilities to apply their knowledge in real-life situations. Studies about students' understanding of hypothesis test have revealed the difficulties that students faced in learning this topic and also how their learning can be enhanced. The next section further discusses these difficulties by categorizing them into the mistakes students make in conducting the hypothesis testing procedure, the misconceptions students have about hypothesis test, and students' incomplete understanding of hypothesis test and the concepts related to it.

## An Overview of Students' Learning Problems in Hypothesis Test

The statistics education scenario at present is such that the aim of teaching and learning of statistics is for students to develop their conceptual understanding of statistics instead of mastering the procedural skills without understanding the reasons behind a certain procedure. However, in learning hypothesis testing, besides having difficulties in understanding the concepts of hypothesis test, it was also found that students encountered difficulties in applying the procedures correctly. This section presents an overview of both the procedural and conceptual difficulties encountered by students in their learning of hypothesis tests.

### Mistakes in the Multi-Step Procedure

According to Smith (2008), students make mistakes at almost every step of the multi-step hypothesis testing procedure. Link (2002) feels that the type of mistakes made by students serves as an indicator of their different levels of understanding. Table 1 shows the categories of mistakes for each hypothesis testing procedure as described by Link (2002) largely. For example, in stating the null and alternative hypothesis, students may have written the wrong population parameter or used the sample statistic instead of the population parameter, and in calculating the value of the test statistic, they may have used the wrong formula or used the correct formula but made a calculation error.

Correctly stating the null hypothesis and the alternative hypothesis is crucial, Link (2002) opines, because the subsequent steps are based on the correctness of this step. For instance, if a student attempted to test for a population mean instead of the population proportion, the student will subsequently use the wrong formula to calculate the test statistic which in turn will affect his or her decision making. However, having stated the correct hypotheses statements does not necessarily guarantee that a student will arrive at the correct hypothesis decision. Students can make mistakes in other steps, for example in forming the probability statement due to calculation mistake or mistake in communicating the decision in the context of the actual problem because they may not know the actual meaning of the results of the hypothesis test (Haller & Krauss 2002) even if they are able to make a valid hypothesis decision in the preceding step.

### Misconceptions about Hypothesis Testing

According to Sotos et al. (2007), students' misconceptions about hypothesis tests have been discussed for more than two decades. They classified students' major misconceptions into misconceptions concerning the different approaches to hypothesis testing, misconceptions concerning the definition of the hypotheses, misconceptions concerning the conditional nature of significance levels, misconceptions concerning the interpretation of p-value, misconceptions concerning the nature of hypotheses tests and misconceptions concerning the statistical significance. Misconceptions concerning the approaches to hypothesis testing are further classified into misconceptions about the philosophy of hypothesis test, and misconceptions that happen in the process of interpreting concepts and results. Meanwhile, misconceptions concerning the definition of the hypotheses refer to students' misunderstanding of the stages in hypothesis testing which was discussed earlier.

TABLE 1. Categories of mistakes in the multi-step procedure of hypothesis testing

| Steps in hypothesis testing | Categories of mistakes |
|---|---|
| Stating the null hypothesis and the alternate hypothesis | 1. stating the wrong population parameter<br>2. stating the sample statistic instead of the population parameter<br>3. use of wrong sign or inequality symbol<br>4. wrong hypothesized value<br>5. meaningless statement |
| Calculating the value of the test statistic | 1. wrong formula<br>2. incorrect calculation<br>3. meaningless response |
| Determining the regions of rejection from the level of significance | 1. failure to distinguish the difference between the one-tailed and two-tailed tests<br>2. incorrect z-value<br>3. meaningless response |
| Generating the critical value and forming the probabilities for the decision criteria | 1. incorrect critical value<br>2. incorrect probability statement<br>3. meaningless response |
| Making a valid decision | 1. wrong decision<br>2. incomplete decision<br>3. correct decision but based on wrong calculations<br>4. meaningless statement |
| Communicating the decision in the context of the problem | 1. superficial or incomplete communication<br>2. no connection to the actual problem<br>3. meaningless statement |

Misconceptions concerning the p-value and misconceptions concerning the statistical significance were identified by Gliner et al. (2002) as two common misconceptions related to the null hypothesis statistical test. In investigating how these misconceptions are treated in textbooks used in education research methods and statistics classes, they found that the textbooks did not relate the concepts of confidence interval and effect sizes to hypothesis testing. Sotos et al. (2009) identified students' level of confidence in their misconceptions of understanding hypothesis test, interpreting p-value and interpreting the significance level. The study involving mostly first year university undergraduates showed a significantly lower level of confidence in interpreting significance level. In addition, students' disinterest in statistics, students' lack of insight and teacher's misconceptions have also been recognized as reasons for students' misconceptions about hypothesis test (Haller & Krauss 2002).

## INCOMPLETE UNDERSTANDING

Students' incomplete or inaccurate or lack of understanding of hypothesis test have been mentioned in literature concerning students' learning of hypothesis test (Link 2002; Lipson et al. 2003). For example, Link (2002) argues that when students use the wrong population parameter in stating the null hypothesis, it is an indication of students' incomplete understanding of the hypothesis testing procedure. Students' incomplete understanding occur both at the procedural and conceptual levels (Evangelista & Hemenway 2002). Although students can successfully execute the procedural steps through memorization and mastery of skills through repetitive exercises, they lack understanding of the underlying concepts. Moreover, in many instances, students focus on the formalism and symbolism instead of on the reasoning of the formalism (Evangelista & Hemenway 2002).

Glaser (2003) found that one of the main reasons for students' incomplete understanding was that they are not taught to make connections between the different concepts of inferential statistics. As such, the students were not able to relate the many inter-related concepts of hypothesis test such as sampling distribution and the significance level. For example, Lipson et al. (2003) found that students view sampling distribution as separated from hypothesis test and also from the sample results. Although it can be time-consuming, students' understanding can be improved by providing different but related problems instead of working with similar situations because as Rossman and Chance (2004) stated, the actual test for students' understanding takes place when students attempt to apply what they have learned. Evangelista and Hemenway (2002) also feel that getting the students to communicate their knowledge to others is a way to reinforce the students' understanding.

## Conclusion

Smith (2008) found that past studies show that students faced many difficulties in developing a deep and connected understanding of statistical hypothesis testing and its' use. She proposed that the understanding of statistical hypothesis test be described from the perspective of the theories of understanding that besides addressing the procedural ability also incorporates the knowledge of the *why*'s and the *how*'s of the procedural steps specifically when applied in contexts. Learning of hypothesis test should focuses on enabling students to use hypothesis test as an inferential tool. The instructional materials should enable students to build understanding of hypothesis test concepts and their ability to extend these concepts to other problem situations. At present there is a need to increase the number of literature in the area of students' learning and understanding of hypothesis because researchers such as Smith (2008) and Sotos et al. (2009) find that there is lack of literature in this area.

## References

Evangelista, F. & Hemenway, C. 2002. The use of the Jigsaw in hypothesis testing. *2nd International Conference on the Teaching of Mathematics at the Graduate Level*. Hersonissos, Crete, Greece.

Garfield, J.B. & Ben-Zvi, D. 2008. *Developing students' statistical reasoning: Connecting research and teaching practice*. Emeryville, CA: Springer.

Glaser, R. 2003. Assessing expert knowledge representations of introductory statistics. *CSE Tech. Rep. No. 600*. Los Angeles. University of California, Center for Research on Evaluation, Standards, and Student Testing.

Gliner, J.A., Leech, N.L. & Morgan, G.A. 2002. Problems with null hypothesis significance testing (NHST): What do the textbooks say? *The Journal of Experimental Education* 71(1): 83-92.

Haller, H. & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research* 7(1): 1-20.

Link, W.C. 2002. *An examination of student mistakes in setting up hypothesis testing problems*. Louisiana-Mississippi Section of the Mathematical Association of America.

Lipson, K., Kokonis, S. & Francis, G. 2006. Developing a computer interaction to enhance student understanding in statistical inference. *Proceedings of the 7th International Conference on Teaching Statistics, ICOTS-7*.

Lipson, K., Kokonis, S. & Francis, G. 2003. Investigation of students' experiences with a web-based computer simulation. *Proceedings of the 2003 IASE Satellite Conference on Statistics and the Internet*. Berlin.

Rossman, A.J. & Chance, B. 2004. Anticipating and addressing student misconceptions. *ARTIST Conference on assessment in Statistics, Lawrence University*. August, 1-4.

Smith, T.M. 2008. *An investigation into student understanding of statistical hypothesis testing*. Doctoral Dissertation. Retrieved from Digital Repository at the University of Maryland. (umi-umd-5658.pdf)

Sotos, C., Vanhoof, S., Noortgate, W. & Onghena, P. 2009. How confident are students in their misconceptions about hypothesis tests? *Journal of Statistics Education* 17(2).

Sotos, C., Vanhoof, S., Noortgate, W. & Onghena, P. 2007. Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review* 2(2): 98-113.

Weinberg, A., Wiesner, E. & Pfaff, T.J. 2010. Using informal inferential reasoning to develop formal concepts: Analyzing an activity. *Journal of Statistics Education* 18(2).

Zieffler, A., Garfiled, J., DelMas, R. & Reading, C. 2008. A Framework to support research on Informal Inferential Reasoning. *Statistics Education Research Journal* 7(2), 40-58.

Saras Krishnan*
Institute of Graduate Studies
Universiti Malaya
50603 Kuala Lumpur
Malaysia

Noraini binti Idris
Faculty of Education
Universiti Malaya
50603 Kuala Lumpur
Malaysia

*Corresponding author: saraskrishnan@yahoo.com

Received: 1 December 2014
Accepted: 29 May 2015