International Journal of Pure and Applied Mathematics
Volume 118 No. 24 2018
ISSN: 1314-3395 (on-line version)
url: http://www.acadpubl.eu/hub/
Special Issue

http://www.acadpubl.eu/hub/

Malaysian College Students Misconceptions in Inferential Statistics

Department of Educational Science, Mathematics and Creative Multimedia, Uni

March 29, 2018

Abstract

Conventional teaching that focuses on computation and remembering procedures is unable to guide students to think statistically. Hence, the purpose of this study is to identify the common misconceptions made by college students in a few topics of inferential statistics. Sixty-eight diploma students who studied in a college were taken as the samples for this qualitative research. The instrument of study was the assessment on statistical reasoning, which consisted of five questions, and was validated by experts. The students were given the opportunity to solve problems by providing answers with reasonable justification through the use of subjective questions. In order to search for the misconceptions, it is necessary for this study to have a variety of students justifications and answers. The overall results show that some of the students are poor in statistical reasoning skills. Furthermore, several misconceptions have been identified through examining the students answers script. The consequence from this finding is important, where the misconceptions could be utilized as a reference for academicians. It might help them to create an effective teaching strategy to minimize the students misconception in college level.

***Key Words*:**Misconceptions; Inferential statistics; College students.

1 Introduction

In Malaysia, the topic of inferential statistics is not just taught in tertiary level, but it has also been included from the beginning of secondary education level. However, this early expose has less effect on the students if the teaching mode still focuses on computation and remembering procedures (1). In addition, most of the statistics assessments are traditional assessments, such as multiple-choice questions and paper-and-pencil tasks (2). These do not encourage the students to think statistically and cause them to possess numerous misconceptions in statistical reasoning as proven in the study of Chan, Ismail, & Sumintono (3). For instance, the common misconceptions in inferential statistics are hypothesis testing, estimation, and so forth (4). Garca-Alonso and Garca-Cruz (5) defined inferential statistics as making inferences about the population through the observation of a sample. From the sample that has been studied, it can enable the researcher to make a generalization in the population of the study.

The challenging part of teaching statistics is on inferential statistics (6). Statistics does not make sense to students if it is taught out of context. Statistics cannot be taught like mathematics, or like physics, as claimed by (7). Generally, teachers must admit that some students do not truly learn statistics until they start to analyze data from their own research. This is one of the problems that make a teachers task to teach the statistical concept become tougher. At the same time, students try to fulfil the needs in acquiring the concept that is given to them before they truly understand their objective on collecting data. This part of statistics usually concerns finding something about a population from a sample taken. From this part, it requires the method of estimating the property of a population through the basis of the sample(8). A good understanding of inferential statistics requires incorporating concepts of sample representativeness and sampling variability to reason the population parameters (5). The objective of this study is to identify the common misconceptions made by Malaysian college students in inferential statistics.

2 Literature Review

Inferential statistics, in the simplest possible terms, is the process to assess the strength of evidence, which concerns a set of observations consistently with a particular hypothesized mechanism that could have been produced by observations (6). The application and interpretation procedure is an essential tool in managing the data analysis in the various fields of research. If the students have problems on the basic nature of statistics, then there is clearly a need to focus more on basics until they have acquired a stable foundation of the discipline (9). As reviewed by (10), there is a common inferential tool that is used to test, which is covered by the statistical hypotheses and computation of P-values, interval estimation, and point estimates. Jala & Reston (10) believed that statistics has many different schools of logical thought that unlike pure mathematics, follow a rigid set. Therefore, it is important to stick to one main approach for inferential statistics.

The ideas of statistical inference have been included in the school curriculum from various countries and are commonly taught by mathematics teachers (6). Because of the relevance and importance of statistical inference, the foundation of statistical course has been in included in high schools. Batanero, Burrill (6) reviewed the classical inferential statistics, which consists primarily of two types of procedures: hypothesis testing and confidence intervals. These techniques are built up based on a scheme of interrelated concepts, which involve probability, random sampling, parameter, distribution of values of a sample statistic, confidence, null and alternative hypothesis, p-value, significance level, and the logic of inference (11). Lessons learned from the researchers in the teaching of statistics reveal much about the variety of obstacles faced by lecturers and students, such as statistics anxiety, negative attitudes, deficits in mathematics backgrounds, and expectations (9).

Consequently, inferential statistics comprises three separate interactions of fundamental elements: (a) the process to reason; (b) the concepts; and (c) the procedural of computations (6). Computations are often easily learnt by students since it can be simulated through the various statistics software. Some of the calculations can be taught by using free software such as Open Office. However, teachers must teach statistics that is incorporated by the three components as mentioned by Batanero, Burrill (6). The teaching method should not only focus on the mechanics of inference, but should also concern the reasoning process (6). Statistical reasoning may be defined as the way people reason with statistical ideas and make sense of statistical information (12). However, statistical reasoning is different from the term of statistical literacy. Statistical literacy includes basic and important skills that may be used in understanding statistical information or research results (12). Research studies on statistical reasoning are still growing and are just beginning to suggest ways to help students develop these outcomes.

In most scientific disciplines, a statistical method is comprehensively used in (a) the description of variability and summarization of empirical results obtained from experimental or observational studies, and in (b) statistical inference based on the observed data (11). This involves an approach when statistical reasoning constructs the idea of which combines data and chance. This leads to the process of interpreting the statistical results. It became apparent that when most statistics educators or researchers assess statistical reasoning, thinking, or literacy, many use different definitions and understandings of these cognitive processes (13). The fundamental idea of this reasoning is a conceptual understanding of important ideas, such as distribution, sampling, uncertainty, and randomness (14). It is also suggested that using real datasets can motivate students to engage in activities, especially when asking them to make conjectures about a dataset before the process of analysis is performed (15).

In the context of real life, statistical reasoning provides advantages when reacting in focusing various types of jobs. For example, it includes psychologists who use statistical information to make judgments and decisions. Statistical information also helps doctors who need to interpret risks and test results. Journalists also use statistical information in the media to make criticism. Political analysts study the trends of their survey in order to interpret polls and elections. Without proper understanding on statistical reasoning, this information has a big possibility to skew towards a dishonest conclusion (13).

3 Methodology/Materials

In order to find the common misconceptions of students, the design of this research is developed by using a qualitative approach. This is because the qualitative method will allow the process to review and analyze the results from students handwritten answers. The result is found through the students solution based on the items constructed by the researcher as the instrument for data collection. The instrument is designed mainly to seek the students misconceptions. It will provide a variety of solutions from the students feedback, which is the main source on how the students solved the statistics problem. Furthermore, the population of the study is students who are taking statistics course, which has inferential statistics as one of its topics. From this population, it may be difficult to identify the students misconceptions because the population is normally represented by a large quantity of data. The researcher has decided to select 68 students randomly who are currently studying in Kolej Profesional MARA Bandar Penawar to be the sample for this researcher. In order to provide data for this study, one set of written test has been designed in order to assess the students reasoning of statistical inference. The item of the test has been inspired by specific guidelines within a discipline to appear more useful than the general categories as stated in Blooms Taxonomy. This guideline focuses on statistical literacy, reasoning, and thinking, which describes that there is an overlap with the types of hierarchy (16). This instrument was constructed with five questions according to delMass (2002) framework to assess the reasoning of students. Example of item was shown in Figure 1. The topic also satisfies to cover inferential statistics in the college syllabus, i.e. estimation and hypothesis testing. Based on the five items constructed, the researcher would like to see how the students will solve this problem and provide their statistical reasoning. The five items have also been validated by two experts before they were distributed to students for the pilot test. Both experts have experience in teaching statistics for Degree and Master levels. Additionally, the questions have been validated by an experienced English lecturer in terms of its use of grammar and language.

The manager of Penawar Chemical Company claims that their major product contains on average μ = 3.0 fluid ounces of acid material per gallon. They further state that the distribution of acid material is normal with a standard deviation of σ = 1.8 fluid ounces. If there is too much acid material, the product will be too dangerous. A quality inspector is brought in to test the product. She randomly selects a sample of 100, gallon-size containers of the product and finds that their mean weight of acid materials is 3.5 fluid ounces per gallon.

a. What are the null and alternative hypotheses that the inspector is exploring? Write them mathematically and explain your reasoning.

b. Conduct the hypothesis test using test statistics and state your conclusion. (Use 5% significant level)

c. Find the probability of getting a sample mean, as high as or higher than 3.5 if the production process was working as normal. Based on this probability would you be inclined to accept or reject the null hypothesis? What conclusion would you draw about the amount of acid material being produced? Explain briefly.

Figure 1. Example of item 4

It is a requirement for the students in this sample to finish all their statistical classes, especially for the topics of normal distribution, estimation theory, and hypothesis testing. At the end of week fourteen during class, the students were told to take this test and needed to prepare their revision if necessary by themselves. During that time, the students have been informed by the researcher regarding the rules of the test, which is necessary as a guideline and reminder to be train the students on the terms of the test, similar to taking an actual examination. The set of items consists of five questions that have been printed confidentially. It was the only instrument used by the researcher and have been administered to 68 students randomly. In other words, it only focuses on students who take statistics class. The test has been conducted in the Seminar Room, which is capable of holding 68 students. Before the test started, the researcher gave a short briefing on the rules of the test to the students in order to ensure the process of this test worked smoothly. After the test was finished, the answer booklets were collected by the researcher himself. The booklets will be examined by the researcher as the important source for the analysis process. This enables the researcher to find the results on students misconceptions.

As mentioned earlier, the aim of this research is to find out the common misconceptions of college students in statistical inference. Hence, the students answer booklets were the only source that the researcher needs to examine manually. The researcher needs to quote the misconceptions if found from the answer booklets written manually by the students. The students answer booklets are the only source of evidence in finding the misconceptions. The analysis conducted would also like to see how the students will respond their answers and provide justifications based on their understanding. The results from the booklets will then be explained clearly on why their answers are classified as misconception, and the justification about this misconception will be highlighted. The descriptive analysis used frequency counts for each part of the questions and percentages of the correct answers. A frequency table has been created by using Microsoft Excel to organize and describe the percentage of correct answers and vice versa. Therefore, it will be easy for the analysis to assist the researcher in writing a summary of the findings, which will be further presented in the next chapter accordingly.

4 Findings and Discussions

According to the evaluation of the results, it can be concluded that some of the students in this study have misconceptions. In item 1, it assessed on how students draw a conclusion from a specific p-value given. However, 4% of the students have been identified to have misconceptions on the role of p-value when conducting the hypothesis. It caused them to put a wrong decision about the null hypothesis. This 4% of students also provided the same faulty conclusion through the question in part (b). This item successfully assessed on how students could respond in giving their conclusion about the null hypothesis from the p-value correctly. 4% of the students have misconceptions on the role of p-value when conducting the hypothesis. One of the misconceptions found is that some students stated that if the p-value is higher, then the null hypothesis is accepted. That student who had this view would appear not to have an understanding about the p-value. They tend to remember the process on what the rule has been said, which causes them to misapply to write the conclusion properly. This tendency is also similar to the observations from the previous study by Chance, delMas, and (13). This is similar with the question in part (1c), which has been designed to assess on how students give reasoning and provide justification about the p-value. It shows that some students have misconceptions when they wrote p-value as a description about an interval. A majority of 62% students have wrote that p-value=0.075 means 92.5% of the samples should give results, which fall in this interval. Other misconceptions that were found included some students interpreting p-value as the probability that the null hypothesis is true. This is one of the misconceptions that has also been observed by (17). Some of them had justified that the higher p-value would mean the null hypothesis would be likely true. (4) also mentioned about the same misconceptions on the p-value interpretations. Concerning the interpretation of p-value, students seem so intuitive about the use of conditional probability in p-value and this has resulted them to face misconceptions (18). It was a serious case when the findings by (17) showed that students believed that lower p-values indicated of having stronger treatment than those with higher p-values. This idea should be avoided because p-value acts as the role in making a decision (16). This is (16) claimed that the p-value is seen as helping to determine which of the hypotheses seem more likely. The null hypothesis is never disproved but may be considered as highly unlikely. Fortunately, only nine students seemed able to understand the interpretation of p-value as they were able to answer question 1 correctly. They were not just able to understand the procedure in conducting the hypothesis testing, but also able to understand the concept of p-value. The item in part c proved how these nine students were able to respond correctly when they encountered the statement, which stated that "p-value=0.075 means that the null hypothesis is rejected at a reasonable alpha. Most of them disagreed with this statement. They were able to provide a correct justification to their answer. Without knowing the rule and procedure properly, there is a possibility for the students to give their reasoning (13). These findings also show that almost 88% of the students know the rule of p-value, but a few of them never understood the role of p-value itself. If a student knows the idea about the true nature of p-value, it would seem unlikely that this misconception can be made. It is important to know that p-value plays an important role whenever a null hypothesis is evaluated because it is important to understand that p-value is a conditional probability and to be aware of what is given (4).

For items 2 and 5, they assessed the students understanding on the topic of estimation and confidence interval. In item 2, it specifically wanted to assess the students reasoning on the effect size of the sample in the confidence interval. 10% students had misconceptions, in which they believed that a large sample will cause a large confidence interval to be set up. This is similar with the misconception found by (8) which identified that the width confidence interval increases with the sample size. These findings were also found by (19), where on their study showed that both university students and professors also had misconceptions on the direct or inverse relations between the interval width, the sample size, and the confidence level. However, in this current study, it is shown that 90% of the students still believed that by increasing the sample of the size, the width of confidence interval will also shrink. Some of them clarified their justification by giving an explanation through the confidence interval formula. In item 5, almost 87% of the students had misconceptions about confidence interval. They seemed to interpret the interval as 95% of the data were included in the confidence interval. Most students mistakenly interpreted that by expecting 95% of the population, it will lie in the interval. Fidler(10) also reported that students typically believed that a given parameter is contained in a confidence interval with a known probability. This misconception about confidence intervals is also similar to the findings from (16). The previous study from (20) also characterized the essence of the error as placing confidence in the interval rather than in the process (Foster, 2014). It also similar with the findings of (21), who confessed through their research that several researchers of other disciplines faced misunderstandings of the interpretation of confidence interval. It is suggested to think that if the experiment was repeated many times and a confidence interval was calculated for each, in the long run, 95% of the intervals would include the population mean (21). This is the fundamental and correct way on how the confidence interval can be understood.

The next items 3 and 4 focused on the finding misconceptions concerning the topic of a hypothesis test. Item 3 showed that most students have a problem in explaining the types of errors, which can possibly occur through the context of given problem. This also shows that most students lack the understanding to apply the rule of conditional probability and its inverse. It also required students to try to provide reasoning for the different types of errors in the context of the problem. Most students failed to discuss the argumentation process, which required them to find the chance of making these errors small. It is different when the idea of alpha as the chance of making a Type 1 error is compared to the idea and role of p-value (16). Both the p-value and terms used as statistically significant are determined to win an argument when it proves a strong and compelling evidence. However, winning an argument by presenting strong evidence may result in the existence of an error when the claim being disputed is actually true (16).

It is really important for students to be able to create null and alternate hypotheses. However, the findings from item 4 show that almost 41% of the students have a comprehension problem to create the first step of defining the null and alternate hypotheses. They were found to be unable to identify the most appropriate statement for each case. The problem in item 4 also needed the students to clearly identify which variable is more important to be considered on creating the hypothesis. The students were faced with a conflict to choose which sign should be put in the alternate hypothesis as indicated in Figure 2. They were shown to provide an incorrect choice of the hypothesis. This similar to the findings of (22), who claimed that sometimes students were confused on how to define the null and alternate hypotheses for a specific context of the problem.

a) **H0:** $\mu = 3.0$
H0: $\mu \neq 3.0$

The null hypothesis state that the major product contains on average $\mu = 3.0$ fluid ounces of acid material per gallon. The alternative hypothesis state that the product contains on $\mu \neq 3.0$ fluid ounces of acid material per gallon.

Figure 2: Student who was not able to write alternate hypothesis correctly

For item 4, it is difficult to know whether the students really understood each element of the stage in hypothesis testing. Based on their written test, it shows that most of them were able to write the procedure of hypothesis testing accordingly. However, this study found that several students had an error in the computation of a critical value in a normal distribution graph. From this error, the students tended to have the most serious misconception that obstructs the understanding of the process of hypothesis testing, which causes the interpretation of its results (23). Although the students might be able to perform all necessary manipulations and formal statistical calculations concerning the hypothesis tests, they have been proved to make misconceptions on the meaning, use, and main concepts of this technique.

5 Conclusion

The findings gave focus on exploring the college students misconceptions in inferential statistics. The overall results showed that some of the students were poor in statistical reasoning skills. They failed to provide justifications against their answers. Not only that, they also harboured misconceptions in the estimation and hypothesis testing; for example, if the p-value is greater than the significance level, there is sufficient evidence to support the null hypothesis and set the confidence level higher, so that the confidence interval will become larger. In other words, college students still have errors in solving the problems of inferential statistics.

Many statistics courses have put statistical inference at the last part of the syllabus. This may encourage informal ideas to be revisited again and again in the class. It may help the students to build their idea to be more intuitive and easier to understand. The topic of statistical hypotheses seems quite difficult as the students need to create arguments. It not a fault when the teaching of statistics is emphasized on solving a series of the textbook problem. However, the educators need to know that learning statistics should be embedded with developing students statistical thinking. The students must be equipped to be able to respond with a correct justification when dealing with a situation where they need to make an inference about data. The educators have the responsibility to determine which of the alternative ways is possible so that the students would be able to understand the concepts behind their work. It is possible to help the students to develop greater understanding than achieving the learning through deductive teaching. Moreover, it is suggested to utilize a model called Statistical Reasoning Learning Environment (SRLE) to develop students statistical reasoning ability, as proposed by (24). It involves six instructional principles, namely focusing on enhancing the conceptual understanding of the basic statistical concepts, utilizing real and motivating data, employing classroom activities, incorporating the use of technological tools, encouraging classroom discourse, and exploiting alternative assessment. The earlier studies conducted by (25), and demonstrated that the SRLE model had a great impact in promoting students statistical reasoning.

References

[1] Chan SW, Ismail Z, Sumintono B. A Refined Technology-Based Statistical Reasoning Assessment Tool in Descriptive Statistics. Advanced Science Letters. 2015;21(7):2352-5.

[2] Chan SW, Ismail Z, Sumintono B. Assessing statistical reasoning in descriptive statistics: A qualitative meta-analysis. Jurnal Teknologi. 2015;72(2):1-6.

[3] Chan SW, Ismail Z, Sumintono B. A Framework for Assessing High School Students' Statistical Reasoning. PloS one. 2016;11(11):e0163846.

[4] Huck SW. Statistical Misconceptions: Classic Edition: Routledge; 2015.

[5] Garca-Alonso I, Garca-Cruz JA. Statistical inference in textbooks: Mathematical and everyday contexts. Proceedings of PME 31. 2007;2:257-64.

[6] Batanero C, Burrill G, Reading C. Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education. 2011.

[7] Ramsey JB. Why do students find statistics so difficult. Proccedings of the 52th Session of the ISI Helsinki. 1999:10-8.

[8] Fidler F. From statistical significance to effect estimation 2005.

[9] Ruggeri K. The Impact of Misunderstanding the Nature of Statistics. Psychology Teaching Review. 2011;17(1):35-40.

[10] Jala LL, Reston E. Graduate Students Conceptions of Statistical Inference. Liceo Journal Higher Education Research. 2013;7(1).

[11] Lr E, editor Statistics: reasoning on uncertainty, and the insignificance of testing null. Annales Zoologici Fennici; 2009: BioOne.

[12] Ben-Zvi D, Garfield J. Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. The challenge of developing statistical literacy, reasoning, and thinking. 2004:3-15.

[13] Garfield JB. Assessing statistical reasoning. Statistics Education Research Journal. 2003;2(1):22-38.

[14] Garfield J, BenZvi D. Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. Teaching Statistics. 2009;31(3):72-7.

[15] Ben-Zvi D. Statistical reasoning learning environment. Em Teia— Revista de Educao Matemtica e Tecnolgica Iberoamericana-ISSN: 2177-9309. 2011;2(2).

[16] Garfield J, Ben-Zvi D. Developing students statistical reasoning: Connecting research and teaching practice: Springer Science & Business Media; 2008.

[17] Gliner JA, Leech NL, Morgan GA. Problems with null hypothesis significance testing (NHST): what do the textbooks say? The Journal of Experimental Education. 2002;71(1):83-92.

[18] Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. Annu Rev Psychol. 2008;59:537-63.

[19] Canal GY, Gutirrez RB, editors. The confidence intervals: A difficult matter, even for experts. Proceedings of the 8th International Conference on Teaching Statistics, ICOTS; 2010.

[20] RobisonCox JF. Having a ball with confidence intervals. Teaching Statistics. 1999;21(3):81-3.

[21] Cumming G, Finch S. Inference by eye: confidence intervals and how to read pictures of data. American Psychologist. 2005;60(2):170.

[22] Batanero C. Controversies around the role of statistical tests in experimental research. Mathematical thinking and learning. 2000;2(1-2):75-97.

[23] Sotos AEC, Vanhoof S, Van den Noortgate W, Onghena P. Students misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. Educational Research Review. 2007;2(2):98-113.

[24] Cobb P, McClain K. Principles of instructional design for supporting the development of students statistical reasoning. The challenge of developing statistical literacy, reasoning, and thinking. 2004:375-96.

[25] Wei Chan S, Ismail Z, Sumintono B. The Impact of Statistical Reasoning Learning Environment: A Rasch Analysis. Advanced Science Letters. 2015;21(5):1211-5.