ASSESSMENT OF GRADUATE STUDENTS' CONCEPTION OF STATISTICAL INFERENCE: PHILIPPINE PERSPECTIVE

Liza Lorena Jala and Enriqueta Reston
University of Cebu, Philippines
University of San Carlos, Philippines
azilanerol@yahoo.com

In the Philippines, graduate education programs in non-statistical disciplines require statistics and research methodology courses, which aim to equip graduate students with the knowledge and skills needed for the applications of statistical tools particularly in thesis or dissertation work. In this paper, graduate students' conceptions of basic ideas and processes of statistical inference were assessed using pre-post classroom self-assessments, interview, examination of students' work samples, classroom observations, and a modified adaptation of a researcher-made Statistical Literacy Assessment Scale. The results revealed several common misconceptions on the concept of sampling and sample representativeness, parameter estimation and in making inferences based on sample data. Based on these findings, an instructional module anchored on constructivist learning will be developed to deepen graduate students' conceptual understanding of statistical inference.

INTRODUCTION

Recent trends in statistics education research have focused on improving teaching and learning processes in the undergraduate introductory statistics courses and in the integration of statistics as an important component of the pre-college mathematics and science curriculum (Gal & Garfield, 1997; Batanero, Godino & Roa, 2004). These researches support the reform efforts in statistics teaching characterized by (1) a shift from mathematical formula-based approaches to conceptual data-based approaches; (2) focus on the impact of students' attitudes, beliefs and conceptions about statistics in relation to their learning outcomes; (3) use of experiential learning approaches and (4) use of varied assessments to assess students learning of statistics. Further, these trends in the teaching of statistics require teachers to become proficient in developing and choosing appropriate assessment methods that are aligned with instruction, and in communicating assessment results to students (Garfield, 1995).

While statistics education at the pre-college and undergraduate levels is important as it lays the foundation for the development of statistical literacy, there is minimal attention given to the graduate level where many of the graduate programs in non-statistical disciplines require statistics and research methodology courses as basic courses in the curriculum. Since graduate education is research-oriented, competence to undertake independent research in an area of specialization most often require graduate students the use of statistical tools and higher levels of statistical literacy. Further, the diverse backgrounds and professional interest and prior knowledge of graduate students brought about by previous education or experience with statistics may have diverse impact on their attitude and beliefs, conceptions about statistics, and learning outcomes in a graduate statistics course. This scenario suggests different approaches to teaching and learning statistics for adult learners particularly for graduate students who will need statistical tools in their research. A more relevant and responsive statistics education to these professionals must offer meaningful learning opportunities for them to connect their knowledge of statistical procedures and research with other related tasks within their degree program and in their respective work settings.

This paper highlights the role of various forms of assessments of graduate students' attitude towards statistics and conceptions on statistical inference to inform and improve instruction in this area. The findings served as basis for curriculum improvement anchored on the constructivist approach for teaching and learning statistics at the graduate level.

The focus on the teaching of statistical inference at the graduate level is relevant as the concepts and methods of Inferential Statistics play a vital role in designing and interpreting empirical results in any scientific discipline. Further, Sotos, Vanhoof, Noortgate, and Onghena (2007) claimed that the interest on statistical inference arises from three realities: (1) it is relevant for the development of research in all empirical sciences; (2) inference receives special attention in statistical courses from almost all scientific areas, where hypotheses tests and confidence intervals

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

are taught to students as the methods for evaluating scientific hypotheses and (3) inferential ideas seem to be especially sensitive to be misunderstood and students are often prone to fall into deep misconceptions (Daniel, 1998). Among the misconceptions mentioned by Sotos, et al. (2007) are the following: (1) misconceptions concerning the definition of hypotheses tests; (2) misconceptions concerning the different approaches to hypotheses testing, (3) misconception on the conditional nature of significance level, which results to the switching of the two terms - significance level and the *p* -value in the conditional probabilities; and (4) misconception when evaluating hypotheses tests that stems from lack of understanding in the difference between statistical and practical significance.

Aside from these misconceptions, it was also found out that many graduate students do not use inferential methods appropriately. Butler (1998) suggested that in spite of the increasing numbers of adults who complete graduate statistics courses, these adults do not often use inferential methods in their research and, when they do try, the results are a shambles. This need for better ways to teach and learn statistical inference at the graduate level is also confirmed by Reston (2007) who found out that graduate student's prior knowledge in statistics was relatively substantial in descriptive statistics but very minimal in inferential procedures commonly used in research.

This study explored graduate students' attitudes and beliefs towards statistics, and assessed their conceptions on statistical inference to provide baseline information that served as basis in the development of an instructional guide anchored on the constructivist perspective of learning.

METHODS

This study used the mixed-methods research approach that combines the collection and analysis of quantitative and qualitative data in stages to investigate and assess graduate students' conceptions on statistical inference. Participants of this study include 18 graduate students enrolled in two graduate statistics courses of a private sectarian university in Cebu City, Philippines. Of these students, seven were enrolled in a doctoral statistics course and eleven were enrolled in the Master's level. Of these, six were males and twelve were females. As to their undergraduate background, they come from various fields such as engineering, education, commerce and philosophy. Most of them are teaching while some are working in offices such as banks and some others are administrators of schools.

At the beginning of the semester, the Survey of Attitudes Toward Statistics–28 (SATS-28 ©) by Schau (2003) and a modified adaptation of a researcher-made Statistical Literacy Assessment Scale (SLAS) by Reston (2004) were administered to the study participants to assess students' attitudes towards statistics and statistical literacy, particularly on interpreting statements involving statistical inference. The SATS-28 © comprises 28 items which assess four components of students' attitudes toward statistics; namely affect, cognitive competence, value and difficulty. The instrument used a 7-point Likert type response scale with higher scores corresponding to more positive attitudes. On the other hand, the researcher-modified SLAS consists of 15 items designed to assess students' statistical literacy in interpreting statements involving statistical inference as used in a given context, and in evaluating generalizations or inferences based on sample data. All items were answerable by "yes", "no", or "cannot tell from the given information" Further, respondents were required to provide a brief explanation or justification for their answers. For students' conceptual understanding of statistical inference upon which misconceptions were derived, sources of data include pre-post classroom assessments, interview with the students, and direct classroom observations.

FINDINGS

Graduate Students' Attitudes and Beliefs towards Statistics

A number of studies gave attention to the assessment and role of beliefs and attitudes in learning statistics (Gal, Ginsburg & Schau, 1997; Schau, Dauphinee, & Del Vecchio, 1992; Carmona, 2002). In the Affective subscale, where students' feelings towards statistics were assessed, two items had a modal score of 7, indicating a high level of agreement to the following statements:

- *I am under stress during statistics class. (Item 14)*
- *I am scared by statistics. (Item 21)*

These responses were validated in the students' interview where some graduate students claim that, in their previous statistics courses they were made to memorize formulas and used those formulas during examination with large hypothetical data sets given to them for statistical analysis.
In the Cognitive Competence subscale of SATS© 28, students' attitudes about their intellectual knowledge and skills when applied to statistics were assessed. Students' responses indicated that most of them neither agree nor disagree about how knowledgeable they are and how skillful they are in terms of statistics application; moreover, most students strongly agree that they can learn statistics. In two of the six items in the subscale, the modal response of 4.00 in the following items indicates students' uncertainty of their cognitive competence in statistics.

- *I have trouble understanding statistics. (Item 3)*
- *I can make a lot of math errors in statistics. (Item 20).*

In the Value subscale, students' attitudes about the usefulness, relevance, and worth of statistics in personal and professional life were assessed. Their responses on the SATS © 28 revealed that most of the graduate students were in total agreement (mode = 7.00) about the usefulness, relevance and worth of statistics in personal and professional life indicating a positive valuing of statistics in their lives. They strongly agree (mode = 7.00) that statistics is worthy and relevant in their lives and can be applied in their profession. They also agreed that statistics skills are needed for them to be more employable and should be part of their professional training.

In the Difficulty subscale, students' attitudes about the difficulty of statistics as a subject were assessed. The graduate students' strongly agreed (mode = 7.00) that learning statistics requires a great deal of discipline and that people have to learn a new way of thinking to do statistics. Moreover, in the interview, the belief that statistics is primarily a mathematical subject dominates and they relate their difficulty to the computational approaches they previously experienced in the course, as evident in the following responses:

- *"I am poor in math, so understandably I am also slow in statistics."*
- *"It was more of computation and memorizing of formulas that I do not remember any of those now."*

Summing up, for most of students, their previous experiences with statistics were primarily computational just like any mathematics subject they had and this was reflected in their attitude towards the course. This is consistent with the findings of Howley (2008) who claimed that the teaching of statistics which focuses on the mathematical rigor does not do any good to students who are not mathematically inclined but leaves the students a bad attitude towards statistics- 'that statistics is really mathematics'.

Graduate Students' Conceptions of Statistical Inference

Results of the Revised SLAS given at the start of the semester revealed graduate students' previous knowledge on statistics and ability to reason with data involving sampling and statistical inferences. In sampling and sample representativeness, students' responses on following item in the SLAS as shown in Figure 1 revealed several misconceptions evident in their statistical reasoning.

In an article entitled "*SWS: 44% Feel Worse Off, Optimism Low*" printed in the Philippine Star (Apr. 20, 2004), the Social Weather Stations (SWS) claimed that a national sample of 1,400 registered voters was *statistically representative* of the Philippine population, with 300 respondents from Metro Manila, 250 from Luzon North of Metro Manila, 250 from Luzon South of Metro Manila, 300 in the Visayas and 300 in Mindanao. Do you agree with this claim? Explain or justify your answer.

Figure 1. Revised SLAS Item 3

Almost all of the graduate students in the study did not agree with the claim. Among the reasons given were as follows:

- *"Samples were just very few, too small, and not enough to represent the 80 million Philippine population"*
- *"Not evenly distributed on the three big regions, Luzon, Visayas and Mindanao."*
- *"Why so many in Luzon area, Luzon is not Philippines, anyway."*

These comments reflect that students' judgments on the representativeness of the sample were made without knowing what sampling procedure was used in the survey. Further, they viewed that the bigger the sample size, the more representative is the sample. Most students concluded that the sample was not representative on the basis of sample sizes in relation to the Philippine population. In sample representativeness, it is important to understand that when the process of selecting the sample has been performed properly, the sample will often have characteristics similar to those of the population (Batanero et al, 1994; cited in Sotos et al, 2007). Understanding these concepts is important in making valid inferences or conclusions about the data.

Classroom observations also revealed that some graduate students have misconceptions on concepts of sampling. One graduate student said *"random sampling is outright sampling"*. The student's understanding of random sampling is simply choosing the respondents according to the researchers' volition. Another surprising revelation related to sampling was a graduate student's comment "*statistics can always be a scientific process in cutting down the number of population into sample*". This comment is a reflection of a graduate student's narrow understanding of statistics which may be attributed to how and what activities they were exposed to, which relates to their previous comments on what activities and how their teacher presents statistics to them.

Another source of misconceptions of many graduate students is the topic on hypothesis testing. It is a major statistical procedure for validating inferences made on populations based on sample data. One major source of misconception on the use of hypothesis testing is evident in many of students' reasoning in their responses to the items in the SLAS where they are not able to distinguish generalizations about populations and statement of facts obtained from samples. This is illustrated in one of the items in the SLAS as shown in Figure 2.

In an article "A land of paradoxes: obesity in the sea of poverty," (Khaleej Times Online, 2004), the author mentioned a study released by the government Food and Nutrition Research Institute and 14 of the country's medical specialty associations which established that obesity is fast consuming a growing number of Filipinos with 52 per cent of women and 12.1 per cent of men found obese among a sample of 4,700 respondents in year 2003. Do these data support the generalization that obesity is associated with the female gender? Explain or justify your answer.

Figure 2. Revised SLAS Item 12

In this item, most graduate students agreed with the generalization that obesity is more associated with females than with males because *"the number speaks the truth"*. This kind of response revealed students' lack of understanding on the concept of statistical inference, that the validity of a generalization cannot be established by just looking at figures from sample data. There is an evident lack of consciousness in their responses on the differences between statements about the sample and generalizations about populations, on the role of chance and uncertainty in making generalizations about populations or about relationships between variables given some sample information. There is need for inferential reasoning and for students' awareness in the necessity to perform an inferential procedure, such as hypothesis testing, prior to accepting generalizations made out of sample data.

Another misconception uncovered among graduate students is on the logic behind inferential reasoning. This is evident in students' responses to the item (see Figure 3) in the Revised SLAS.

In an experimental study on the effectiveness of a computer-assisted language program at private sectarian university, the researchers used two groups of college students enrolled in a basic English course. The experimental group was taught using the innovative language program while the control group was taught using conventional classroom procedures. Both groups were given pretests and posttests. As part of the findings, the researchers reported in a local university journal the following: "*Considering the difference between the Pre-test and Post-test means of the control group which is 18.5 compared to that of the experimental group which is 11.94, it may be deduced arithmetically that more learning happened in the control group*." Do you agree with the researchers' report? Explain or justify your answer.

Figure 3. Revised SLAS Item 12

In this item, most graduate students agreed with the researchers' report by simply comparing figures obtained from the samples. One student reasoned out that *"numbers won't lie and since it was shown based on arithmetic deduction"*. This kind of reasoning indicates a lack of understanding that conclusions in research based on sample information are statistically inferred and not "arithmetically deduced."

These misconceptions may be indicative of poor undergraduate background on inferential statistics. To address these misconceptions, different teaching strategies were used in these graduate statistics classes. These include the critique of newspaper articles or researches published in journals and magazines which deal with application of statistics. Graduate students were exposed to these kinds of activities so they would understand better the key concepts in statistics, especially on statistical inference.

CONCLUSIONS AND RECOMMENDATIONS

It is evident from the findings of this study that graduate students' attitudes and beliefs about statistics were greatly influenced by their previous experiences in learning statistics in their undergraduate or even secondary education. Their misconceptions on the concepts of sampling, sample representativeness, and the logic of inferential reasoning when dealing with sample data also stemmed from poor background knowledge on inferential statistics in their undergraduate statistics courses. Assessing students' attitudes and conceptions on statistical inference at the graduate level provided baseline information on their misconceptions which needed to be addressed in their graduate statistics courses. This provided the basis for interventions needed to improve graduate students' conceptual understanding and reasoning about statistical inference. To address these misconceptions, a curricular guide anchored on the principles of constructivism was developed and pilot tested in another graduate statistics course. This curricular guide will provide a mixture of activities and exposition leading students to explore statistical inference concepts and construct their own conceptual understanding of important concepts in inferential statistics.

In view of these findings, it is recommended that pre-assessments of students' prior knowledge, attitudes and beliefs should be made an integral part of assessment in graduate statistics courses as basis for more appropriate pedagogical approaches in teaching statistics to graduate students with diverse backgrounds. Also, assessments should emphasize conceptual understanding rather than mastery of computations and must be based on contextualized activities that require students to apply their understanding in the management, analysis, evaluation and communication of real data arising from different research contexts.

REFERENCES

Batanero, C., Godino, J., & Roa, R. (2004). Training Teachers to Teach Probability. *Journal of Statistics Education*, *12*(1).

Butler, R. S. (1998). On the Failure of the Widespread Use of Statistics. *Amstat News*, *251*(84).

Carmona, J. (2002). La teoría de facetas y el escalamiento multidimensional en la elaboración y validación de cuestionarios de actitudes: Una aplicación al dominio de las actitudes hacia la estadística. Unpublished doctoral thesis, Universidad de Sevilla.

Daniel, L. G. (1998). Statistical significance testing: A historical overview of misuse and misinterpretation with implications for the editorial policies of educational journals. *Research in the Schools*, *5*(2), 23–32.

Gal, I., & Garfield, J. B. (1997). Curricular Goals and Assessment Challenge in Statistics Education. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*. Amsterdam: IOS Press and International Statistical Institute.

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring Attitudes and Beliefs in Statistics Education. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*, Amsterdam: IOS Press and International Statistical Institute.

Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review*, *63*(1), 25-34

Ottaviani, M. G., & Batanero, C. (1999). The Role of the IASE in Developing Statistical Education*. In 6th Islamic Countries Conference on Statistical Sciences Proceedings.*

Reston, E. (2004).The Impact of Computing Technology on Teaching and Research: Implications for Statistical Education. *University of San Carlos Graduate Journal*. *21*(1), 24-31

Schau, C. (2003). *Students' attitudes: The "other" important outcome in statistics education*. Paper presented at the Joint Statistical Meetings, San Francisco, CA.

Schau, C. G., Dauphinee, T., & Del Vecchio, A. (1992). The development of the Survey of Attitudes Toward Statistics. *American Educational Research Association*, San Francisco, CA.

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.