ASSESSMENT OF TERTIARY STUDENTS' LEARNING OF STATISTICAL MODELLING

Vaughan van Appel
University of Johannesburg
vvanappel@uj.ac.za

Rina Durandt
University of Johannesburg
rdurandt@uj.ac.za

ABSTRACT – In this paper, we report on the learning of statistical modelling in a second-year statistics module through the assessment of a problem that required a Monte Carlo simulation. On the forefront of the 4th industrial revolution is science, technology, engineering and mathematics subjects, where mathematical statistics plays a key role in topics such as machine learning and predictive analysis. Students often find statistical modelling difficult, where obstructions in the modelling process could lead towards a dead-end. For this reason, assessment of learning, for learning and as learning in statistics education seems necessary. General pillars of good assessment practice is considered in this study, as well as guidelines for the development of students' conceptual understanding of the content, such as, statistical reasoning, statistical thinking and statistical literacy. Therefore, this study was conducted to provide educators with information of student achievement of desired student learning outcomes. Based on an analysis of the reports collected individually and through voluntary group work, descriptive statistics are presented. These results are discussed in relation with assessment measures and provides a basis for teaching and learning statistics.

Keywords: Assessment; Statistics Education; Statistical Modelling; Monte Carlo Simulations; Tertiary Education

INTRODUCTION

The demands of the 4th industrial revolution (4IR) and the adequate preparation of students for these demands is a recent and central topic of discussion at institutions of higher learning, and particularly in South Africa (e.g., Hussin, 2018; Xing & Marwala, 2017,). On the forefront of the 4IR discussion is science, technology, engineering and mathematics subjects (compare Idin, 2018). More specifically, mathematical statistics and particularly statistical modelling plays a key role in 4IR, where topics such as machine learning and predictive analysis, to name just a few, is necessary. Educating students in programming languages to be able to unite man and machine is fast becoming a necessary skill in statistics modules, but also in other science and engineering related fields. Statistics and statistical modelling originates from the subject mathematics and more precisely mathematical modelling, where both students and educators find the latter challenging (compare Blum & Borromeo Ferri, 2009). Garfield and Franklin (2011) argue that the difference between mathematicians and statisticians is related to how they view and assess data analysis. Statisticians view data analysis as a *process* that involves formulating a scientific question that can be answered with data, followed by designing a plan to collect the data, then collecting and analysing the data with appropriate techniques, and finally interpreting the results as they relate to the original scientific questions. One such procedure is to educate tertiary statistics students in programming languages and to act as true statisticians by using Monte Carlo simulations (named after the city Monte Carlo in Monaco). This kind of simulation of a real-life problem situation is different from a physical experiment as it performs repeated random sampling of the experiment on a computer program to obtain numerical results.

According to the National Research Council (1989, p. 69) it is important for educators to consider "What is tested is what gets taught. Tests must measure what is most important." Following this notion, Garfield and Franklin (2011) argue that the three pillars of assessment, namely, *cognition*, *observation* and *interpretation* should carefully be considered in the planning of learning activities and the assessment of such activities. Furthermore, students find the learning of statistics difficult and their attitudinal scores can decrease over time (compare Van Appel & Durandt, 2018). Researchers in statistics education suggested

cognitive statistical learning outcomes related to statistical literacy, statistical reasoning and statistical thinking (compare Garfield & Ben-Zvi, 2007).

The broad purpose of this study was to measure statistics students' learning of statistical modelling through the assessment of a planned learning activity based on a Monte Carlo simulation. Furthermore, through achieving the desired learning outcomes, students should become better equipped for the demand of 4IR. The idea in this study is to gather meaningful information about the students' learning of statistical modelling, and to better align teaching and assessments in a second-year statistics course. Formally, the research question is: *To what extend have students demonstrated cognitive statistical knowledge (literacy, reasoning and thinking) when attempting the Monte Carlo simulation learning activity in a programming language.*

THEORETICAL PERSPECTIVES

In assessing students' statistical learning, educators should consider the foundational pillars of good assessment practice (Pellegrino, Chudowsky & Glaser, 2001), as well as the guidelines for the development of students' conceptual understanding of the statistical content and the desired learning outcomes (compare Garfield & Ben-Zvi, 2007), and the criteria for suitable learning activities (Garfield & Franklin, 2011). This research initiative was grounded in a pragmatic view (Creswell, 2013) and we carefully considered a combination of the following three theoretical perspectives.

The first perspective relates to the three foundational pillars, namely, *cognition*, *observation*, and *interpretation*, which encompass an "assessment triangle" (Pellegrino et al., 2001). These pillars should ideally form the foundation of all assessment practices. Following this notion, Garfield and Franklin (2011) explained the purpose of assessment is connected to all three pillars and assessment practices should be *of* learning (more summative oriented), *for* learning (more formative oriented by providing feedback to students) and *as* learning (oriented as a combination between the summative and the formative placing the student central between learning and assessment). The latter can gestalt in statistics courses through examples that ask from students to create a unique model in an authentic activity where they have the opportunity to reflect and make sense of their own knowledge throughout the creation process.

The second perspective, widely supported by researchers in statistics education (e.g., Garfield & Franklin, 2011), informed this inquiry regarding the categorisation of cognitive statistical learning outcomes (Garfield & Ben-Zvi, 2007):

i. *Statistical literary* – understanding and using the basic language and tools of statistics.
ii. *Statistical reasoning* – reasoning with statistical ideas and making sense of statistical information.
iii. *Statistical thinking* – recognising the importance of examining and trying to explain variability and knowing where the data came from, as well as connecting data analysis to the larger context of a statistical investigation.

The third theoretical perspective relates to the viewpoint from Garfield and Franklin (2011) that informed the selection of the statistical modelling learning activity in this inquiry, which considers the role of cognition by a set of guiding principles. Some of these principles are: to include real data and real problem context, to include recognising and understanding the concept of variability, to include opportunities to select methods of graphing and analysing data, to maintain a balance between items assessing, understanding probability concepts and understanding statistics concepts, and when is it appropriate to require students to provide interpretations of data analysis as well as justifications for their analyses and conclusions.

As a combination, the abovementioned notions informed this inquiry in the selection of the learning activity, in the specification of the learning outcomes and in the assessment of students' individual and group activity sheets. More specifically, our intention was to balance procedural proficiency, conceptual understanding and the use of context of a statistical investigation through assessment of the learning activity.

RESEARCH DESIGN AND METHODOLOGY

Participants

This inquiry involved a sample of 118 second-year statistics students studying towards a BSc degree in Mathematical Science or a BSc degree in Actuarial Science at a large public university in Johannesburg. All students enrolled in this module are majoring in the mathematical sciences, passed their first-year mathematics and statistics courses, and performed above average in high school mathematics.

Monte Carlo simulations and the statistical modelling learning activity

A Monte Carlo simulation is widely regarded as a very useful approach in solving complex applications in statistics (e.g., Zickar, 2005). More specifically, Monte Carlo simulations are computer driven simulations of the problem using known prior information or parameters to generate plausible random sample data. Thereafter, the generated data are used to evaluate statistics of interest, for example likelihoods, expected values and variability (compare Mooney, 1997; Paxton et al., 2001; Ross, 2013). In addition, Monte Carlo simulation allows one to visualise the potential outcomes of the experiment, which may aid in better overall decision making. In order to conduct such simulated experiments, students need to be educated in using a computer programming language to analyse and solve real world problems (this is commonly known as Education 4.0). Normal practice for students studying towards a degree in Mathematical Sciences is to expose them early on in their undergraduate statistics module to a programming language – for example, Excel in their first year of study and *R* in their second year of study.

This statistical modelling learning activity was a simplified real world problem and students were required to follow the steps of a Monte Carlo simulation and compile a short scientific report on their findings. In short, we expected the students to conclude whether the potential reward is worth the risk or not. At this stage of their professional development, statistics students have already received the required exposure to complete the learning activity successfully. Formal exposure to statistical content is through theory lectures (where theoretical content is introduced), tutorial sessions (where theoretical problems are solved) and practical classes (where students are exposed to real world problems and introduced to the programming language *R*). Therefore, the students were expected to solve the learning activity by using the programming language *R* in which they have already received the appropriate preparation. Figure 1 displays the modelling activity used in this inquiry (Braun & Murdoch, 2007, p. 110) and Figure 2 shows an example solution. Students' were given the activity during the last week of the module and were given the choice to complete the task individually or in groups (maximum of three learners per group).

Simulate the following simple model of auto insurance claims:

- *Claims arise according to a Poisson process at a rate of 100 per year.*
- *Each claim is a random size following a gamma distribution with shape and rate parameters equal to 2 and 4, respectively. This distribution has a mean of R0.5 million and a variance of R0.125 million^2.*
- *The insurance company must pay claims on the day they arise.*
- *The insurance company earns premiums at a rate of R53 million per year, spread evenly over the year (i.e. at time t measured in years, the total premium received is 53t.)*

Write R code to do the following:

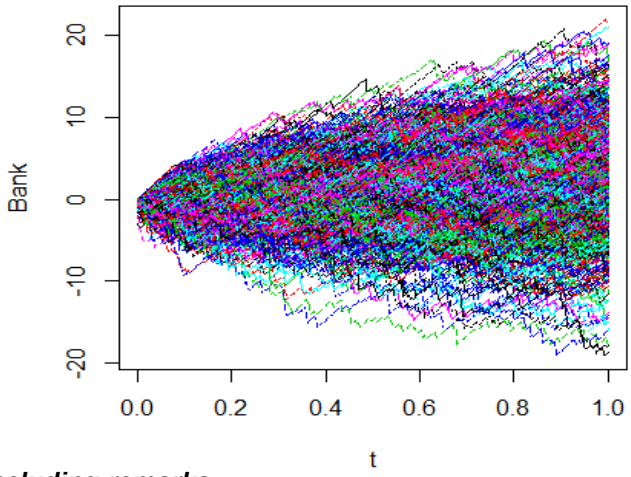
a) *Simulate the times and amounts of all the claims that would occur in one year. Draw a graph of the total amount of money that the insurance company would have through the year, starting from zero: it should increase smoothly with the premiums, and drop at each claim time.*
b) *Repeat the simulation 1000 times, and estimate the following quantities:*
    i. *The expected minimum and maximum amount of money that the insurance company would have at t=1.*
    ii. *The expected final amount of money that the insurance company would have at t=1.*
    iii. *Comment on the total amount of money that the insurance company would have through the year.*

c) *Carry out any further calculations to enable you to decide whether this is a good business model or not? State whether you would be interested in investing in this insurance company or not. Give a reason(s) for your answer.*

Figure 1: The statistical modelling learning activity (source Braun & Murdoch, 2007, p. 110)

Example solution

- *The expected minimum and maximum amount of money that the insurance company would have at t=1 is -23.04436 and 21.0134 million Rand respectively.*
- *The expected final amount of money that the insurance company would have at t=1 is 3.088 million Rand.*
- *At many instances, the total money that the insurance company has is negative. This implies that the insurance company will need to have access to some credit or finance facility to be able to settle all claims.*



Example Concluding remarks

- *It is not a good business model, since the probability that the bank balance stays positive throughout the full year is only 7.9%. Therefore, it is very likely that a finance facility would be needed, which would drastically reduce the potential profit, or*
- *the probability that the portfolio will have a positive balance at year-end is 71.3%. Therefore, I would say it is a good business model.*

Figure 2: Example solution of the statistical modelling activity

The example solution (see Figure 2) contains a description of some of the information that was expected in the students' reports, and therefore was used as a guideline in the assessment process, integrated with the notion form Pellegrino et al. (2001) and Garfield and Franklin (2011). For example, students needed to display an understanding of risk in their concluding remarks. That is, we all have different risk tolerances and students should make the choice of whether the risk is worth the reward or not. This free-response item (at the end of the activity) allowed the students to explain and communicate their understanding.

STATISTICAL ANALYSIS AND DISCUSSION OF RESULTS

Student answers from the statistical reports were marked and categorised according to the three proposed categories that originated from the literature framework – statistical literacy, statistical reasoning, and statistical thinking (Garfield & Ben-Zvi, 2007). A fourth category was added based on the guidelines from Garfield and Franklin (2011) as it seemed appropriate for students at second-year level to submit a statistical report that showed the processes followed, suitable graphs with descriptions, and concluding remarks as an interpretation of their findings. The grades were allocated according to the scheme: 0 – poor, 1 – somewhat satisfied, and 2 – satisfied. A mark out of two was awarded for each category. Afterwards, a specialist in the field of mathematics education checked all grades. Table 1 displays the descriptive statistics generated from the 37 group reports received. A holistic view of the findings shows that many students struggled to: (i) understand how to combine all the information given in the real-life

problem to correctly answer the problem on their own; (ii) implement the problem in *R*; and (iii) construct a neat concise statistical report (report appearance). We expected a better-quality solution from students and were concerned about the students that could not even make sense of the data to start with the first step of the Monte Carlo simulation.

Table 1: Descriptive Statistics

	Literacy	Reasoning	Thinking	Report appearance
Mean	1.24	0.70	0.62	0.97
Median	2.00	0.00	0.00	1.00
Mode	2.00	0.00	0.00	1.00
Standard Deviation	0.86	0.85	0.79	0.80
Kurtosis	-1.48	-1.31	-0.91	-1.41
Skewness	-0.51	0.63	0.81	0.05
Minimum	0.00	0.00	0.00	0.00
Maximum	2.00	2.00	2.00	2.00

From Table 1, the statistical reasoning (M = 0.70), thinking (M = 0.62) and report appearance (M = 0.97) received below satisfactory results on average. More specifically, most individuals/groups received a poor result for their reasoning and thinking, and a somewhat satisfactory result on their report appearance. A crucial point to make is the low quality of the statistical reports – students could not correctly express the statistics and had no idea how to compile or present their findings in a neat and well-structured report. Report writing is a necessary skill for a professional statistician and should require more attention during their formal professional development.

We were also interested in the students' answers of the free-response item at the end of the activity sheet showing how they displayed and communicated an understanding of risk. The responses of four respondents are shown below:

Respondent 1: *"We would invest in the insurance company since the amount is dependent on the claim."*

Respondent 2: "This could be a good business model because both parties have mutual relationship."

Respondent 3: "Yes, we are interested in investing in this company because even though there are drops in the curve, the curve still increases. Therefore, we will make money on our investment. The biggest loss the company can experience is expected to be R3.03m and the most profit is expected to be R6.08m".

Respondent 4: "No, we will not choose to invest in this company. According to our calculations only 66.98% of the 1000 simulations have positive cash flows by the end of the year. Hence the project is only profitable 66.98% of the time. This is therefore a very high risk company to invest in. Also, the variance from the 1000 simulations was R37.51 million with a standard deviation of R6.12 million. Since the mean is only R3.33 million the standard deviation is much bigger than the mean which indicates very volatile cash flows. Hence, profitability of the investment is very unpredictable."

Respondents 1 and 2 are both examples of poor responses, as they both provided no meaningful information to answer the problem and show no sign of statistical reasoning or thinking in their responses. Respondent 3, is a somewhat satisfactory response that shows an understandable reasoning in the conclusion, however, the answer is lacking further statistical information (which was alluded to in point (c) of the modelling activity) to construct an improved conclusion. In addition, there was not a strong presence of the risk factor in their conclusion. Thus, the response is lacking in the statistical thinking category. Respondent 4 is an example of a satisfactory response. More specifically, this respondent clearly highlighted the potential risk by calculating the likelihood of a positive balance and the potential spread of the profit at year-end, showing statistical reasoning and thinking.

CONCLUSION

The broad purpose of this inquiry was to measure statistics students' learning of statistical modelling through the assessment of a planned learning activity based on a Monte Carlo simulation. Furthermore, through achieving the desired learning outcomes, students could become more prepared for the demand of 4IR. The main idea was to gather meaningful information about the students' cognitive statistical knowledge, and the evaluation of students' learning to inform a more integrated and balanced assessment practice versus teaching approach. General pillars of good assessment practice have been considered, as well as guidelines for the development of students' conceptual understanding of the content, such as statistical reasoning, statistical thinking, and statistical literacy. The research question '*To what extend have students demonstrated cognitive statistical knowledge (literacy, reasoning and thinking) when attempting the Monte Carlo simulation learning activity in using a programming language?'* was answered by interpreting descriptive statistics of the data.

Results revealed some second-year statistics students struggled with organising the data in a statistical modelling learning activity, and more than half of the students performed below average in the statistical reasoning and thinking categories. Although, a few statistical reports were of high quality, the majority lacked important features. With these results, educators in statistics have an improved understanding of students' misconceptions and required skills and it could lead to a more desirable answer of 'what evidence do educators need to show student's understanding' and 'will this assessment provide the evidence'?

REFERENCES

Blum, W., & Borromeo Ferri, R. (2009). Mathematical Modelling: Can it Be Taught and Learnt? *Journal of Mathematical Modelling and Application*, *1*(1), 45–58.

Braun, W. J. & Murdoch D. J. (2007). A First Course in Statistical Programming with R. *Cambridge University Press*.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75(3), 372-396*.

Garfield, J., & Franklin, C. (2011). Assessment of learning, for learning, and as learning in statistics education. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 133-145). Springer, Dordrecht.

Hussin, A. A. (2018). Education 4.0 made simple: Ideas for teaching. *International Journal of Education & Literacy Studies, 6(3), 92-98.*

Idin, S. (2018). An Overview of STEM Education and Industry 4.0, *Research Highlights in STEM Education, 194-208*, ISRES Publishing.

Mooney, C. Z. (1997). *Monte Carlo simulations*, Thousand Oaks, CA: Sage.

National Research Council. (1989). *Everybody counts: A report to the nation o the future of mathematics education. Washington*, DC: National Academy Press.

Paxton, P., Curran, P., Bollen, K.A., Kirby, J. & Chen, F. (2001), Monte Carlo experiments: design and implementation, *Structural equation modeling, 8, 287-312*.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press, 2102 Constitutions Avenue, NW, Lockbox 285, Washington, DC 20055.

Ross, S. M. (2013). *Simulation (5th ed.)*, San Diego, CA: Academic Press.

Sigal, M. J. & Chalmers, R. P. (2016). Play it again: teaching statistics with Monte Carlo simulations, *Journal of Statistics Education, 24 (3), 136-156*, DOI: 10.1080/10691898.2016.1246953.

Van Appel, V. & Durandt, R. (2018). Dissimilarities in attitudes between students in service and mainstream courses towards statistics: an analysis conducted in a developing country*. EURASIA Journal of Mathematics, Science and Technology Education, 2018, 14(8), em1565*. https://doi.org/10.29333/ejmste/91912

Xing, B. & Marwala, T. (2017). Implications of the fourth industrial age on higher education, *The Thinker: For the Thought Leaders (www.thethinker.co.za), 73, 10-15.*

Zickar, M. J. (2005). *The psychology research handbook: A guide for graduate students and research assistants (2nd ed.)*, Thousand oaks, CA: Sage Publications.