13th International Congress on Mathematical Education
Hamburg, 24-31 July 2016

A FRAMEWORK OF PROBABILITY CONCEPTS NEEDED FOR TEACHING REPEATED SAMPLING APPROACHES TO INFERENCE

Hollylynne S. Lee — NC State University, USA
Helen M. Doerr — Syracuse University, USA

In this paper, we describe theoretical and empirical perspectives to articulate what learners should understand about an approach to inference that emphasizes a process of randomizing data, repeating through simulation, and rejecting any model with observed data in the extreme of a distribution. Our work revealed that there were key probability concepts that could assist learners in developing richer understandings and capabilities to a repeated sampling approach to inference. Our perspectives and framework are presented herein.

INTRODUCTION

Repeated sampling approaches to inference have recently become prominent in reform-oriented statistics curricula, often using the power of computing tools for creating and running simulations of a repeated sampling process. Researchers have argued that a simulation that uses repeated sampling can be an important tool to help students develop a deep understanding of the abstract statistical concepts involved in inferential reasoning (Burrill, 2002; Maxara & Biehler, 2006). In 2007, Cobb suggested that educators help students develop an understanding of inference through the "three R's: randomize data, repeat by simulation, and reject any model that puts your data in its tail" (p. 12). In this paper, we aim to unpack this "three R's" approach to illuminate the role of probability concepts in a simulation approach to inference and to offer a framework that captures the key probabilistic conceptions and capabilities that need to be addressed when teaching a repeated sampling approach to inference.

Simulations of repeated sampling have been used in several collegiate curricula efforts in the United States, and researchers have reported modest results in improvement of students' understandings of inference through this approach (e.g., Garfield, delMas, & Zieffler, 2012; Tintle, et. al., 2012). In addition, new curriculum standards in some countries, such as the United States and New Zealand, suggest such an approach for high school students (Council of Chief State School Officers, 2010; New Zealand Ministry of Education, 2006).

Many curriculum developers recommend that learners experience repeated sampling methods in a physical way before using computing power (e.g., Cobb, 2007; Rossman, 2008). The physical experiences are intended to assist learners in conceiving of the process of sampling as a repeatable action (e.g., Watson & Chance, 2012). These physical experiences serve as a way to reveal the underlying probability assumptions in a problem (e.g., is it equally likely for two events to occur? Does each person have an equal chance to being assigned to a treatment group?). However, in many curricula, the processes of the repeated sampling are often created by the instructor (or curriculum), and students are asked to use this prescribed process with physical objects or are told exactly what to input in a computer simulation (e.g., Cumming, Miller, & Pfannkuch, 2014; Roy et al., 2014). Such prescriptions likely serve to mask both the underlying randomization that is taking place and the probability assumptions that are being made. A simulation approach, using physical and computer tools, seems to be an appropriate way to help students develop statistical inference conceptually. Thus, while the "three R's" highlight the important elements of a simulation model, understanding each part of a simulation and the relationships among the parts is conceptually complicated. In fact, students who know *how* to conduct a simulation may not have a robust understanding of *why* they are conducting a simulation, *what* is being simulated, and *how* to make appropriate conclusions based on a simulation.

THEORETICAL PERSPECTIVES

Similar to the work of Borovcnik (2014) and others, we consider learners' use of probability models as essential to conceptualizing a repeated sampling approach to inference. When conducting a simulation, one is trying to model some process so as to potentially better understand the inputs, inner workings, and outputs of the process. At the core of a simulation involving a stochastic process is randomness—the first of the "three R's". However, at every step of a simulation-based repeated sampling approach to inference there are other issues concerning probability models that learners need to understand. What follows are descriptions of the theoretical perspectives we use in our work related to probability models, the models and modeling perspective on learning, and repeated sampling representations and processes described by others.

While probability is an abstract concept that cannot be directly measured, probability is used as an expression of likelihood of an event. There are many objects and real world events whose behavior and outcomes cannot be completely determined ahead of time, even when there is a great deal of information about actions on the object or event in world. Whether we are faced with many or few contraints, we can often build models to express a probability of an event occurring. A relatively simple example is to consider that in a given toss of a number cube, we cannot determine the outcome of which side will land facing up; but, we can build a probability model to estimate the likelihood of each side facing up. How one engages in building such a model may differ based on the approach or perspective taken (e.g., classical, frequentist or subjective) (Borovcnik & Kapadia, 2014).

Chaput, Girard and Henry (2011) described three parts of a probability modeling process that includes translating observations and assumptions of contextual problem into a pseudo-concrete working model, mathematizing the model into a hypothesis-driven probability model that can be enacted, and validating a model through examining how a model fits with empirical data and interpreting the model within the context of the problem. Many have advocated that we want students and teachers to understand the bi-directional relationship between probability models and data, and between empirically-developed models, typically from a frequentist perspective, and theoretically-developed models, typically from a classical perspective (e.g., Eichler & Vogel, 2014; Konold & Kazak, 2008; Lee & Lee, 2009; Pfannkuch & Ziedins, 2014; Pratt, 2011; Stohl & Tarr, 2002; Wild, 2006). We claim that in using a repeated sampling approach to inference, the probability assumptions and the model-building process in a simulation should be made more explicit. Pfannkuch and Ziedins' (2014) description of probability models and their purpose provide a useful perspective on the strong role that probability models have in a repeated sampling approach to inference:

> A probability model will often be associated with the idea of a system evolving dynamically over time…a model is usually built to answer a particular question or questions about a system, sometimes just to understand its behavior better, but often in order to optimize some measure of its performance, or alternatively, to predict performance under some alternative scenario. …although they [models] are only approximations to what happens in the real world, these approximations can help us better understand the behavior in the real world (p. 103).

In a repeated sampling approach to inference, learners are building and using models as approximations to what happens in the real world under conditions of randomness. To begin with, learners should be conceiving of observed outcome(s) from an observational study or an experiment as resulting from a process that is repeatable, and that repeating the process may result in a different outcome. This is a frequentist perspective of probability. A key question becomes, how unusual is what happened in the particular instance that we just observed? In other words, what is the likelihood of a particular observed outcome occurring if a process is repeated many times? That is the end goal of a repeated sampling approach to inference. But to achieve this, we need to make sense of the problem we are trying to solve and consider some of the underlying assumptions, what process is being repeated, and what may be the role of randomness and probability in that process. All of these considerations require a model building process.

Given our focus on probability as a model, it made sense to situate our work in a models and modeling perspective on teaching and learning mathematics, as articulated by Lesh and Doerr (2003). Using this perspective, the goal for a learner is to build a model that can be generalized to other situations and productively re-used. Thus, we are particularly interested in how learners can develop a robust model of using repeated sampling for making inferences for problem situations. We beleive such a model includes understanding relationships among a problem situation, physical enactments of sampling, representations of those enactments, computer representations, the underlying randomization (i.e., the probability models discussed above), the distribution of the statistics of interest, and how to interpret and use such a distribution to make a decision about likelihood of an event.

USE OF REPEATED SAMPLING AND SIMULATIONS

Much research and curriculum development in recent years has focused on understanding inference and simulation approaches. For example, Saldanha and Thompson (2002) report that when students can visualize a sampling process through a three-tier scheme, they develop a deeper understanding of the process and logic of inference. This scheme is centered around “the images of repeatedly sampling from a population, recording a statistic, and tracking the accumulation of statistics as they distribute themselves along a range of possibilities” (p. 261). In their work, Saldanha and Thompson expicitly have students experience a three-level sampling process that includes: 1) randomly drawing items to form a sample and record a statistic of interest, 2) repeating this process a large number of times and accumulating a collection of sample statistics, and 3) partitioning the collection of statistics to determine what proportion lies beyond a given value.

Lane-Getaz (2006) also describes the process of using a simulation to develop the logic of inference starting with a question in mind, “what if?”, to investigate a problem including three tiers: population parameters, random samples, and distribution of sample statistics (see Figure 1). In line with Lane-Getaz’s suggestion, Garfield et al. (2012) used a generalized structure of the logic of a simulation approach to inference in their curriculum materials. Their structure includes specifying a model, using the model to generate simulated data for a single trial and then multiple trials, each time collecting a statistic of interest, and finally using the distribution of collected statistics to compare observed data with the behavior of the model.

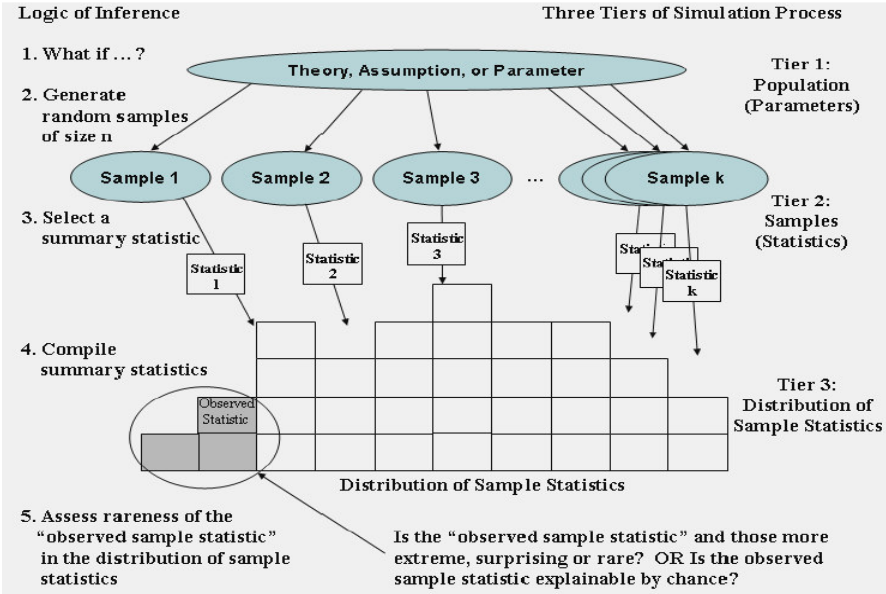

Figure 1: Representation of Lane Getaz‘ simulation process model (2006, p. 280).

In 2014, Saldanha and Liu described work with learners in repeated sampling tasks and made the case that students should develop a stochastic conception of an event that “entails thinking of it as an instantiation of an underlying repeatable process, whereas a non-stochastic conception entails thinking of an event as unrepeatable or never to be repeated” (p. 382). Such a stochastic conception includes seeing an event as an expression of some process that could be repeated under similar conditions that produces a collection of outcomes and “reciprocally, seeing a collection as having been generated by a stochastic process” (p. 382). This approach was also emphasized by Lee, Starling and Gonzalez (2014) in their work using empirical sampling distributions to help learners compare the likelihood of real world events by examining which event occurred less often under a repeated stochastic process.

EMPIRICAL PERSPECTIVES

For the past several years, we have worked together in the design and implementation of a graduate course for those interested in teaching statsitical thinking. Participants in the course tend to be those interested in teaching statsitics in for learners ages 13-20. We used a models and modeling perspective (Lesh & Doerr, 2003) to design a sequence of model development activities consisting of structurally related tasks that begin with a *model eliciting activity,* and are followed by *model exploration activities* and *model application activities* (Arleback, Doerr, & O’Neill, 2013). Our aim was to foster the development of understandings of how probability models are critical for understanding repeated sampling approaches to inference. In our work with teachers as learners, our learning goal is for teachers to develop a stochastic conception of events and a conceptual model that they can use to approach inference situations using a repeated sampling approach, and for them to be able to assist others in using such an approach. This conceputal model includes understanding the relationships among the problem situation, physical enactments of sampling, representations of those enactments, computer representations, and the underlying randomization (i.e., the probability models discussed above), the distribution of the statistics of interest and how to interpret and use such a distribution to make a decision. In order for teachers to develop that model (and the entailments needed for teaching that model to other learners), we hypothesized that the teachers should be able to make connections to and use an underlying probability model of repeatable actions with unpredictable outcomes. Our instructional design and analysis of teachers' conceptualizations is not the focus of this paper, but can be read in other publications (Arnold et al., forthcoming; Lee et al., in press; Lee et al., 2015).

To frame the key probabilistic conceptions for teaching a sampling approach to inference, we drew on empirical data from our research with teachers and the theoretical perspectives represented in literature (e.g., Saldanha & Thompson, 2002; Lane-Getaz, 2006; Garfield et al., 2012). We outline below several important conceptualizations that we believe are powerful part of a learners' conceptual model for a repeated sampling approach to inference.

FRAMEWORK OF PROBABILITY CONCEPTIONS IN INFERENCE

If a goal is to have learners' understand how and why a repeated sampling approach to inference works, more attention needs to be given to the modeling process, the explicit role of probability in inference, and use of probabilistic language. We feel that there is a two-level modeling process that should be made explicit to learners. The first level is the process of creating a local specific model of the real world context in statistical terms. The second level is creating a simulation process that embodies the repeatable actions in the original context and can be used to generate random samples. Most previous theoretical work has combined these two levels into a single "population" level (e.g., see Figure 1). We found in our research that those teachers who could only vaguely state that they needed a model to begin with often had difficulty in analyzing a new context and designing an appropriate way to use a simulation with physical or computer tools. It was the teachers who were able to carefully unpack the assumptions in the context and conceive of what actions are being repeated in the context who were able to apply such understandings to build new models and design appropriate simulations in new contexts.

We also suggest that repeated sampling approaches to inference need to be more explicit about building a distribution of sample statistics, assisting learners in viewing this distribution as a probability distribution, using the distribution to reason about the observed statistic, and making a claim about the chance of an observed statistic (and those more extreme) occurring. Our work has led us to articulate key conceptualizations involved in a repeated sampling approach to inference and the capabilities such conceptions afford (Table 1). These conceptualizations seem important for learners to develop in order to have a robust way of conceiving how a repeated sampling approach using simulations can be used to engage in inference in a range of contexts. We saw evidence of this in our learners who *were* developing stronger conceptions, and hypothesize that those who struggled may have benefitted from learning experiences that would help them develop such conceptualizations. On the right side of Table 1, we identify what each conceptualization can allow learners to enact while they are using a repeated sampling approach to inference.

Relating back to the “Three R’s” approach (randomize, repeat, and reject) suggested by Cobb (2007) and used by many others, our first two conceptualizations clarify what one needs to understand about the first R of randomize. Part of the second and third conceptualizations can help learners understand the meaning of the repeat phase. And finally, the fourth and fifth conceptualizations are critical in understanding decision-making that occurs in the reject phase. We believe that it is crucial for learners to conceive of the distribution of sample statistics as an empirical probability distribution from which the likelihood of events can be evaluated by examining the relative frequency for a range of events to occur.

Conceptualization	**Capabilities this conception affords**
1. Conceive of events in the real world problem as a result from a repeatable action	Identify the underlying probability model of the event of interest (what is repeatable?) Consider what results would be considered unusual, or what would be considered usual or likely to happen. Express a usual expectation as a null hypothesis. Specify the observed statistic and the statistic of interest that should be observed when each action is repeated.
2. Conceive of and create a method for simulating the repeated sampling process	Identify the repeatable action that needs to be enacted Choose tool (physical or computer) and map the action in the real word to a simple repeatable process using the tool.
3. Conceive of repeated sampling as a way to generate simulated statistics	Recognize need to enact process for a random sample of same size n and record statistic for event of interest. Repeat random sample k times (large number) and collect statistics from each sample for event of interest.
4. Conceive of how collected statistics from repeated samples vary with respect to likelihood, and thus a distribution of such statistics can be conceived of as a probability distribution	Build a distribution of the recorded statistics Notice what seems to be usual (typical, or more likely to occur), and what is unusual (or unlikely to occur) Locate the original observed statistic in the distribution, and consider whether it was in a range of “likely to happen” or “unlikely to happen”.
5. Conceive of the inferential decision as deciding if the observed statistic and those more extreme are explainable by chance.	Partition the distribution of recorded statistics using the observed statistic as a partition border. Use proportional reasoning to evaluate the likelihood that the observed statistic, and those more extreme, happened under the random process used to generate repeated actions and simulated statistics.

Table 1: Key conceptions and capabilities for understanding repeated sampling for inference

Because our target learners were those interested in teaching statistics, our focus was on assisting them to understand how probability is used in a repeated sampling approach and to develop a generalized model as a connected conceptual system that they can draw upon as they in turn assist their students in learning to use a repeated sampling approach to inference. It is important to recall that all of our teachers had previous exposure and experience with learning traditional inference techniques, and some had experiences in teaching such techniques. The key conceptualizations we describe seem useful for all learners to develop. In this regard, it is our intent that instructional designers and researchers could use our framework of conceptualizations and what they afford to inform research and teaching focused on developing learners' understandings of repeated sampling approach to inference.

References

Ärlebäck, J. B., Doerr, H. M., & O'Neil, A. H. (2013). A modeling perspective on interpreting rates of change in context. *Mathematical Thinking and Learning*, *15*, 314–336.

Arnold, P., Confrey, J., Jones, S., Lee, H. S., & Pfannkuch, M. (forthcoming). Learning trajectories in statistics education. In D. Ben-Zvi, J. Garfield, & K. Makar (Eds.), *First International Handbook of Statistics Education* (pp. TBD). Springer.

Borovcnik, M. (2014). Modelling and experiments–An interactive approach towards probability and statistics. In T. Wassong, D. Frischemeier, P. R. Fischer, R. Hochmuth, & P. Bender (Eds.) *Mit Werkzeugen Mathematik und Stochastik lernen–Using Tools for Learning Mathematics and Statistics* (pp. 267-281). Springer Fachmedien Wiesbaden.

Borovcnik, M. & Kapadia, R. (2014). A historical and philosophical perspective on probability. In E. J Chernoff & B. Sriraman, (Eds.), *Probabilistic thinking: presenting plural* perspectives (pp. 7-34). New York: Springer.

Chaput, B., Girard, J.-C., & Henry, M. (2011). Frequentist approach: Modelling and simulation in statistics and probability teaching. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 85-95). The Netherlands: Springer.

Cobb, P. (2007). The introductory statistics course: A ptolemaic curriculum? *Technology Innovations in Statistics Education, 1*(1).

Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington D.C.: Author.

Cumming, J., Miller, C., & Pfannkuch, M. (2014). Using bootstrap dynamic visualizations in teaching. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA*. Voorburg: The Netherlands: International Statistical Institute.

Eichler, A., & Vogel, M. (2014). Three approaches for modelling situations with randomness. In E. J. Chernoff & B. Sriraman (Eds.), *Probablistic Thinking* (pp. 75-99). Dordrecht, The Netherlands: Springer.

Garfield, J., delMas, R., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM, 44*, 883-898.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education, 2*(1), 1-37.

Lane-Getaz, S. J. (2006). What is statistical thinking and how is it developed? In G. Burrill (Ed.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp. 272-289). Reston, VA: National Council of Teachers of Mathematics.

Lee, H. S., Angotti, R. L., & Tarr, J. E. (2010). Making comparisons between observed data and expected outcomes: Students' informal hypothesis testing with probability simulation tools. *Statistics Education Research Journal, 9*(1), 68 - 96.

Lee, H. S., Doerr, H. M., Tran, D., & Lovett, J. N. (in press, 2016). The role of probability in developing learners' models of simulation approaches to inference. *Statistics Education Research Journal 15*(2), p. TBD.

Lee, H. S., Starling, T. T., & Gonzalez, M. D. (2014). Using data to motivate the use of empirical sampling distributions. *Mathematics Teacher, 107*(6), 465-469.

Lee, H. S., Tran, D., Nickell, J., & Doerr, H. M. (2015). Simulation approaches for informal inferences: Models to develop understanding. In K. Krainer & Vondorva, N. (Eds.), *Proceedings of the Ninth Congress of European Society for Research in Mathematics Education*. (pp. 707-714). Online: http://1drv.ms/1mncPDq.

Lesh, R. A., & Doerr, H. M. (Eds.). (2003). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Mahwah, NJ: Lawrence Erlbaum Associates

New Zealand Ministry of Education. (2006). *The New Zealand curriculum: Mathematics and statistics. Nzcurriculum.tki.org.nz*

Pfannkuch, M., & Ziedins, I. (2014). A modelling perspective on probability. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking* (pp. 101-116). Dordrecht, The Netherlands: Springer.

Pratt, D. (2011). Re-connecting probability and reasoning about data in secondary school teaching, In *Proceedings of the 58th World Statistics Conference*, Dublin. Retrieved from http://2011.isiproceedings.org/papers/450478.pdf.

Rossman, A. (2008). Reasoning about informal statistical inference: One statistician's view. *Statistics Education Research Journal, 7*(2), 5-19.

Roy, S., Rossman, A., Chance, B., Cobb, G., Vanderstoep, J., Tintle, N., & Swanson, T. (2014). Using simulation/randomization to introduce p-value in week 1. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute.

Saldanha, L., & Thompson, P. W. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics, 51*, 257-270.

Stohl, H., & Tarr, J. E. (2002). Developing the notions on inference using probability simulation tools. *Journal of Mathematical Behavior, 21,* 319-337.

Tintle, N., Topliff, K., Vanderstoep, J., Holmes, V.-L., & Swanson, T. (2012). Retention of statistical concepts in a preliminary randomization-based introductory statistics curriculum. *Statistics Education Research Journal, 11*(1), 21-40.

Watson, J., & Chance, B. (2012). Building intuitions about statistical inference based on resampling. *Australian Senior Mathematics Journal, 26*(1), 6-16.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal, 5*(2), 10-26.