



Simulations of the Sampling Distribution of the Mean Do Not Necessarily Mislead and Can Facilitate Learning

David M. Lane
Rice University

Journal of Statistics Education Volume 23, Number 2 (2015),
www.amstat.org/publications/jse/v23n2/lane.pdf

Copyright © 2015 by David M. Lane, all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

Key Words: Sampling distribution simulations; Variance of means; Effectiveness of simulations; Central Limit theorem.

Abstract

Recently [Watkins, Bargagliotti, and Franklin \(2014\)](#) discovered that simulations of the sampling distribution of the mean can mislead students into concluding that the mean of the sampling distribution of the mean depends on sample size. This potential error arises from the fact that the mean of a *simulated* sampling distribution will tend to be closer to the population mean with large sample sizes than it will with small sample sizes. Although this pattern does not change as a function of the number of samples, the size of the difference between simulated sampling distribution means does and can be made invisible to observers by using a very large number of samples. It is now practical for simulations to use these very large numbers of samples since the speed of computers and even mobile devices is sufficient to simulate a sampling distribution based on 1,000,000 samples in just a few seconds. Research on the effectiveness of sampling distribution simulations is briefly reviewed and it is concluded that they are effective as long as they are used in a pedagogically sound manner.

1. The Problem

Intuitively, it would seem that students would have little difficulty understanding that the mean of the sampling distribution of the mean is the population mean. However, [Watkins, Bargagliotti, and Franklin \(2014\)](#), in a careful and insightful analysis, found that students often mistakenly believe that this mean is affected by sample size. Importantly, these authors showed that this misconception can result from interacting with a simulation of the sampling distribution of the mean. The misconception occurs because the mean of a simulated sampling distribution will not

exactly equal the population mean and, since this mean is based on all values of all samples, it will tend to be closer to the population mean with large samples than with small samples. In a compelling example presented by Watkins et al., students simulated a sampling distribution of the mean using 100 samples for sample sizes of 5, 15, and 30. The means of the resulting three simulated sampling distributions were therefore based on 500, 1,500 and 3,000 values respectively. The differences between the simulated sampling distribution means and the population means decreased as a function of sample size: The absolute differences were 0.092, 0.040, and 0.005 standard deviations for sample sizes of 5, 15, and 30 respectively. The students noticed this pattern in their simulation results and, understandably but incorrectly, concluded that the larger the sample size, the closer the mean of the sampling distribution is to the population mean.

Watkins et al. argued that a common solution to this kind of problem, increasing the number of samples, does not solve the problem because this pattern is not unlikely for any number of simulations. On this basis, Watkins et al. concluded that this problem cannot be fixed.

2. The Problem can be Fixed by Increasing the Number of Samples

The key question from the point of view of teaching statistics is whether increasing the number of samples can make this sample size effect so small as to be invisible to the student. For example, the default distribution in the sampling distribution simulation described in [Lane and Scott \(2000\)](#) has a mean of 16, a standard deviation of 5, and displays the results of simulations to four significant digits that include two decimal places. Therefore, if the mean of a simulated distribution is between 15.995 and 16.005, it will be rounded off to 16.00, thus obscuring the potentially misleading influence of sampling error. In terms of standardized units, this will occur if the absolute difference between the simulated mean and the population mean is less than 0.001 standard deviations.

[Table 1](#) shows the probability that the absolute standardized difference between the mean of a sampling distribution and the population mean will be exceeded by less than 0.001 standard deviations as a function of sample size and the number of samples. As noted, differences less than 0.001 standard deviations will not be apparent when the displayed means are rounded to two decimal places. These calculations were based on the normal distribution and will be accurate for even an extremely non-normal distribution because the distributions of means are based on the product of the number of samples and the sample size, a value that is 500 or more for every entry in the table.

Table 1. Probability of Sampling Error Being Visible after Rounding to Two Decimal Places

Number of Samples	Sample Size			
	5	10	15	20
100	0.9822	0.9748	0.9691	0.9643
100,000	0.4795	0.3173	0.2207	0.1573
1,000,000	0.0253	0.0016	0.0001	0.0000
2,000,000	0.0016	0.0000	0.0000	0.0000

[Table 1](#) demonstrates that sampling error showing up in results displayed to two decimal places does not become very unlikely with less than 1,000,000 samples and, for a sample size of 5, with less than 2,000,000 samples.

Is it practical for a simulation to have as many as 1,000,000 or even 2,000,000 samples? Clearly it is not practical in simulations for which students compute the mean of each sample individually and combine the results to form a distribution. However, it is not impractical when the simulation software does the work. For example, 100,000 samples with a sample size of 20 can be simulated on an iPad Air using a sampling distribution simulation written in JavaScript (available at http://onlinestatbook.com/stat_sim/sampling_dist/index.html) in under a second. I timed a test version that is not yet online and it did 2,000,000 simulations in just under 10 seconds. Naturally, different simulation platforms will differ in their speed.

To be on the safe side, students should always be reminded that a simulated sampling distribution is only an approximation of a sampling distribution and that very small differences between simulation results should not be taken seriously. That, itself, is a good lesson in understanding randomness and reinforces the idea that a sampling distribution is a theoretical distribution and not a frequency distribution.

In addition to having a sufficient number of simulations, it is also important that the pseudo-random number generator perform well. Although developers are often at the mercy of random number generators such as those built into Java or JavaScript on a web browser, they should be sure that the simulation results match the theoretical expectations to the number of decimal places displayed.

3. Are Simulations Effective?

Although not their main focus, Watkins et al. presented a rather negative assessment of the evidence for the effectiveness of sampling distribution simulations. Specifically, they stated “The formal research that has been conducted to compare student understanding of sampling distributions following instruction with and without simulation generally has found no difference or a modest difference in favor of simulation.” Four sources are cited to support this conclusion

([Mills 2002](#); [Meletiou-Mavrotheris 2003](#); [Chance, delMas, and Garfield 2004](#); and [Pfaff and Weinberg 2009](#)). However, a close look at these sources suggests a more positive assessment. Although [Mills \(2002\)](#) stated that there were few studies that empirically evaluated sampling distribution simulations, she did cite two studies ([delMas, Garfield, and Chance 1999](#); [Weir, McManus, and Kiely 1990](#)) that found benefits of simulations, although the Weir et al. study found the benefit only for lower-ability students. Mills did not cite any studies that failed to find that sampling distribution simulations were effective. [Meletiou-Mavrotheris \(2003\)](#) did not compare a simulation to a non-simulation group but was, instead, interested in whether the interactive Fathom-based simulations would be more effective than Minitab's "black box" simulations previously found to be relatively ineffective. Unlike the black box simulations, the interactive simulations led to "fairly sophisticated understandings of sampling distributions and inferential statistics" (p. 290). [Chance et al. \(2004\)](#) concluded that mere exposure to sampling distribution simulations is unlikely to significantly change students' deep understanding, but they were more optimistic that simulations could be effective when implemented well. Based on their research, they provided recommendations for creating conceptually enhanced simulations tools. Finally, [Pfaff and Weinberg \(2009\)](#) studied a card-based simulation of the central limit theorem rather than a computer-based simulation and did not find evidence that the simulation was effective. It is possible that the manual work with the cards interfered conceptual processing.

Although the issue of the effectiveness of sampling distribution simulations is far from settled, a number of studies suggest that they are effective ([delMas et al. 1999](#); [Lane and Tang 2000](#); [Mills 2004](#); [Weir et al. 1990](#); [Ziemer and Lane 2000](#)). The [Mills \(2004\)](#) study is especially notable because she found a very large effect size (Cohen's $d = 1.20$) for improvement from the pre- to post-test in the simulation condition and no evidence for improvement in the non-simulation condition. Finally, in their review of research on teaching and learning statistics, [Garfield and Ben-Zvi \(2007\)](#) reviewed studies of sampling distribution simulations and concluded that simulations "can play a significant role in enhancing students' ability to study random processes and statistical concepts" (p. 9).

4. Using Simulations Effectively

One way to use simulations is to allow students to experiment with a simulation and to discover the important principles on their own. However, this kind of "pure discovery learning" has been shown to be very ineffective ([Mayer 2004](#)). Alternatively, the instructor could take complete control and incorporate a simulation in a lecture without student interaction. Unfortunately, watching a simulation performed by an instructor can lead students to be passive observers and learn very little ([Lane and Peres 2006](#)). A more effective method is to engage students by asking them to predict the results of a simulation and then use the simulation to confirm or disconfirm their prediction ([de Jong, Hartel, Swaak, and van Joolingen 1996](#); [Garfield and Ben-Zvi 2007](#); [delMas et al. 1999](#); [Weiman 2005](#)). [Garfield and Ben-Zvi \(2007\)](#) summarized their findings as follows: "By forcing students to confront the limitations of their knowledge, we have found that students are more apt to correct their misconceptions and to construct more lasting connections with their existing knowledge framework" (p. 312).

Applying this to teaching about the sampling distribution of the mean, students could be asked to predict how the means of simulated sampling distributions will vary as a function of sample size.

Those who predict that sample size will have an effect will see their prediction disconfirmed and can be expected to learn more than if this were simply demonstrated. The fact that the mean of the sampling distribution of the mean equals the population mean irrespective of the population *shape* could be taught in a similar manner.

5. Conclusions

[Watkins et al.'s \(2014\)](#) important observations about simulations of the sampling distribution of the mean should be taken to heart by anyone using or developing such a simulation. Instructors using a simulation of the sampling distribution of the mean should be aware of the way the simulation can be misleading and make sure to take steps to keep students from being misled. In general, using a very large number of samples will be sufficient. Developers of simulations should make sure that enough samples are always taken and/or advise the student about the proper interpretation and the potential to be misled. Research indicates that using a simulation is an effective way to teach about sampling distributions as long as it is used in a pedagogically sound manner.

Acknowledgements

I would like to thank David W. Scott for comments on a previous version of this article.

References

Chance, B., delMas, R., and Garfield, J. (2004), "Reasoning about Sampling Distributions," in *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, eds. D. Ben-Zvi and J. Garfield, Dordrecht: Kluwer, 295–323.

de Jong, T., Hartel, H., Swaak, J., and van Joolingen, W. (1996), "Support for Simulation-based Learning; the Effects of Assignments in Learning about Transmission Lines," In A. Diaz de Ilarazza Sanchez and I. Fernandez de Castro (Eds.), *Computer Aided Learning and Instruction in Science and Engineering*, 9-27. Berlin: Springer Verlag.

delMas, R., Garfield, J., and Chance, B. (1999), "A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning," *Journal of Statistics Education*, 7, <http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm>.

Garfield, J. and Ben-Zvi, D. (2007), "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics," *International Statistical Review*, 75, 372-396.

Lane, D. M. and Peres, S. C. (2006), "Interactive Simulations in the Teaching of Statistics: Promise and Pitfalls," In Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Bahia, Brazil, <http://www.ime.usp.br/~abe/ICOTS7/Proceedings/index.html>

Lane, D. M. and Tang, Z. (2000), "Effectiveness of Simulation Training on Transfer of statistical concepts" *Journal of Educational Computing Research*, 22, 383-396.

Lane, D. M. and Scott, D. W. (2000), "Simulations, Case Studies, and an Online Text: A Web-Based Resource for Teaching, Statistics," *Metrika, special issue on Interactive Statistics*, 51, 67-90.

Mayer, R. E. (2004), "Should there be a three-strikes rule against pure discovery learning?" *American Psychologist*, 59, 14.

Meletiou-Mavrotheris, M. (2003), "Technological Tools in the Introductory Statistics Classroom: Effects on Student Understanding of Inferential Statistics," *International Journal of Computers for Mathematical Learning*, 8, 265-297.

Mills, J. D. (2002), "Using computer simulation methods to teach statistics: A review of the literature," *Journal of Statistics Education* 10,
<http://www.amstat.org/publications/jse/v10n1/mills.html>

Mills, J. D. (2004), "Learning abstract statistics concepts using simulation," *Educational Research Quarterly*, 28, 18-33.

Pfaff, T. J. and Weinberg, A. (2009), "Do hands-on activities increase student understanding?: A case study," *Journal of Statistics Education* 17,
<http://www.amstat.org/publications/jse/v17n3/pfaff.html>

Watkins, A., Bargagliotti, A. E., and Franklin, C. (2014), "Simulation of the Sampling Distribution of the Mean can Mislead," *Journal of Statistics Education*, 22,
<http://www.amstat.org/publications/jse/v22n3/watkins.pdf>

Weir, C. G., McManus, I. C., and Kiely, B. (1990), "Evaluation of the Teaching of Statistical Concepts by Interactive Experience With Monte Carlo Simulations," *British Journal of Educational Psychology*, 61, 240-247.

Weiman, C. (2005), "Science Education in the 21st Century—Using the Tools of Science to Teach Science," Paper presented at the National Science Teachers Association, Dallas, TX, April.

Ziemer, H. and Lane, D. (2000), "Evaluating the Efficacy of the Rice University Virtual Statistics Lab," Poster presented at the 22nd Annual Meeting of the National Institute on the Teaching of Psychology, St. Petersburg Beach, FL, January.

David M. Lane
MS-25
6100 Main Street
Rice University, Houston TX 77005

Lane@rice.edu
Phone: 713-364-5052

[Volume 23 \(2015\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Guidelines for Readers/Data Users](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)