Pedagogical Simulation of Sampling Distributions and the Central Limit Theorem

KEYWORDS:
Teaching;
Sampling Distributions;
Central Limit Theorem;
Simulation;
Visual Basic;
Microsoft Excel.

Reidar Hagtvedt
Georgia Institute of Technology, USA.

Gregory Todd Jones
Georgia State University, USA.

Kari Jones
University of Georgia, USA.
e-mail: lawgtj@langate.gsu.edu

Summary
Students often find the fact that a sample statistic is a random variable very hard to grasp. Even more mysterious is why a sample mean should become ever more Normal as the sample size increases. This simulation tool is meant to illustrate the process, thereby giving students some intuitive grasp of the relationship between a parent population and the distribution of a sample mean.

◆ INTRODUCTION ◆

It is usually not difficult to make the point that the variation of a sample mean of a number of observations of a random variable is less than the variation of a single observation of a random variable. For example, students immediately understand that the mean heights of classes of students will vary less than the mean heights of the individual students themselves. Nevertheless, it is often difficult for the students to remember that each time a sample is collected, the sample mean will take on a new value, according to a distribution that is obviously related to the parent population, but with key differences. That is, they readily accept the fact that the distribution of the sample mean has the same mean as the parent population, but the smaller variance, and especially the bell shape, are ideas that remain somewhat elusive.

The simulation tool that is the subject of this article, Visual Sampling Distributions, was written in Visual Basic for Applications and integrated into a Microsoft Excel spreadsheet in order to provide a widely accessible means for students to witness how multiple samples generate different values for the sample mean, and how that mean is distributed. The hope was that 'seeing is believing'. After the student chooses the parent population distribution, the sample size and the number of samples, the software generates relative frequency histograms to illustrate the distributions.

As an added bonus, these histograms clearly illustrate that the variability of the sample means is dramatically reduced when compared to the variability of the individual observations. While the exact relationship

$$\sigma_{\bar{x}} = \sigma / \sqrt{n}$$

may not be obvious due to the simulation, it does become clear that a larger sample size lowers the standard error.

◆ DESCRIPTION ◆

After an opening welcome page, the student is provided with an introduction to the relevant concepts and complete instructions for use. Then the student chooses simulation parameters (default

© 2007 The Author
Journal compilation © 2007 Teaching Statistics Trust

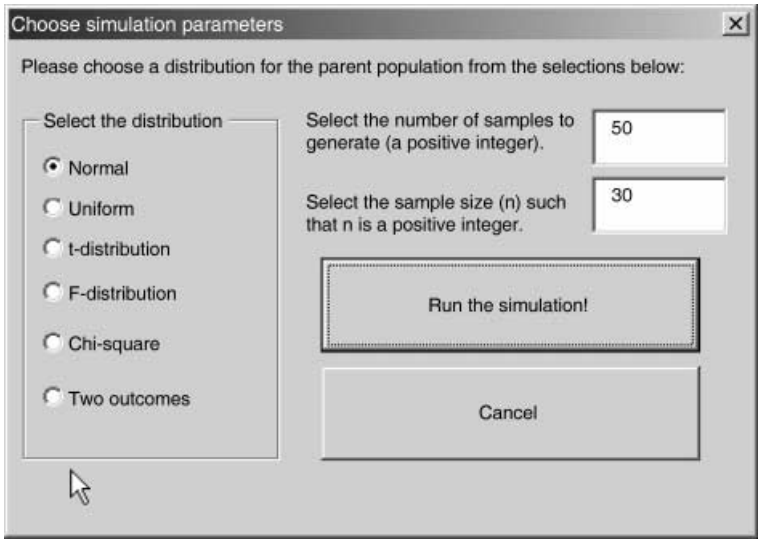

Fig. 1. Simulation parameters

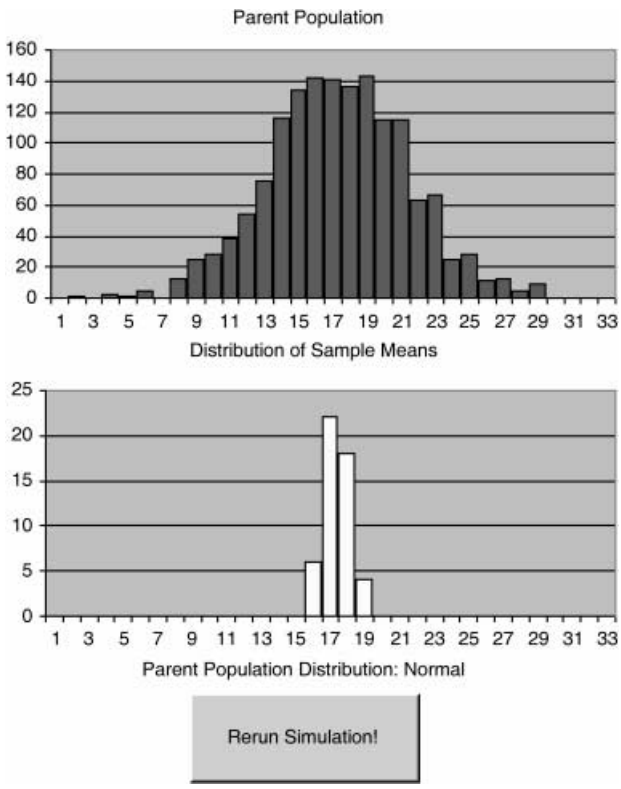

Fig. 2. The histograms

parameters of a Normal parent population distribution with 50 samples of size 30 are provided) (see figure 1).

When the student presses the 'Run the Simulation!' button, the software generates underlying data, calculates sample means and populates two relative frequency tables and their related histograms (see figure 2).

The 'Rerun Simulation!' feature allows for a systematic exploration of the parameter-space. Multiple simulations will hopefully illustrate the following core points:

- With a larger sample size, the variation in the distribution of the sample means becomes smaller.
- Any parent population will yield a sample mean that approaches a Normal distribution if the sample size is sufficiently large, i.e. the Central Limit Theorem.

The relative frequency histogram of the sample mean distribution may not look Normally distributed at all. For example, if a student selects a sufficiently small number of samples, the sampling distribution is swamped by random variation. See figure 3. With a larger number of samples, it is more likely that the distribution of sample means will appear relatively more bell shaped and symmetrical (see figure 4).

◆ EVALUATION ◆

We asked 28 students to evaluate the software, both in writing and with a small quantitative survey, as an extra credit assignment in an introductory statistics course. Another 30 introductory statistics students had access to a different program, but did the same statistical problems in their survey, hence providing a rough control. Some limited guidance was provided to motivate experimentation. We suggested that students begin with a large number of samples, such as 100, in order to ensure that the distribution of sample means was sufficiently populated to be meaningful. We also recommended that students begin with a

© 2007 The Authors
Journal compilation © 2007 Teaching Statistics Trust

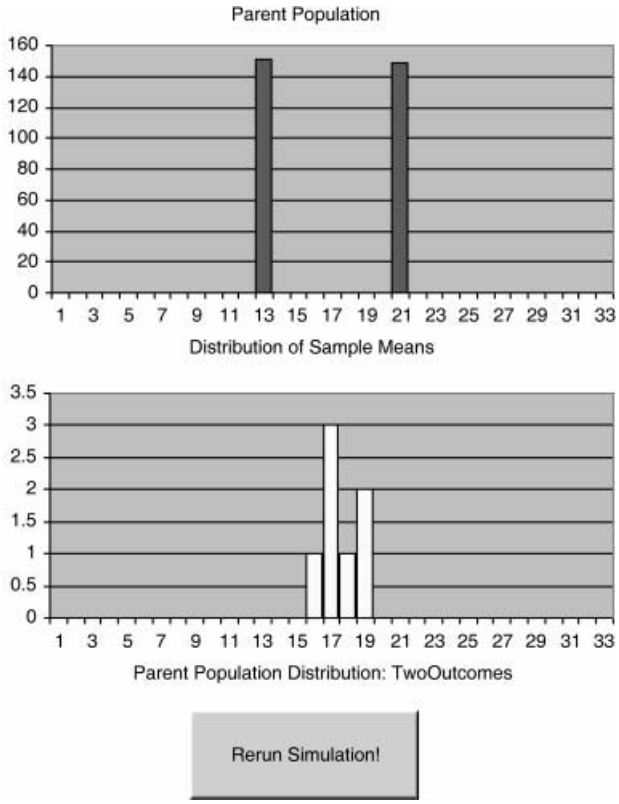

Fig. 3. A small number of samples (10)

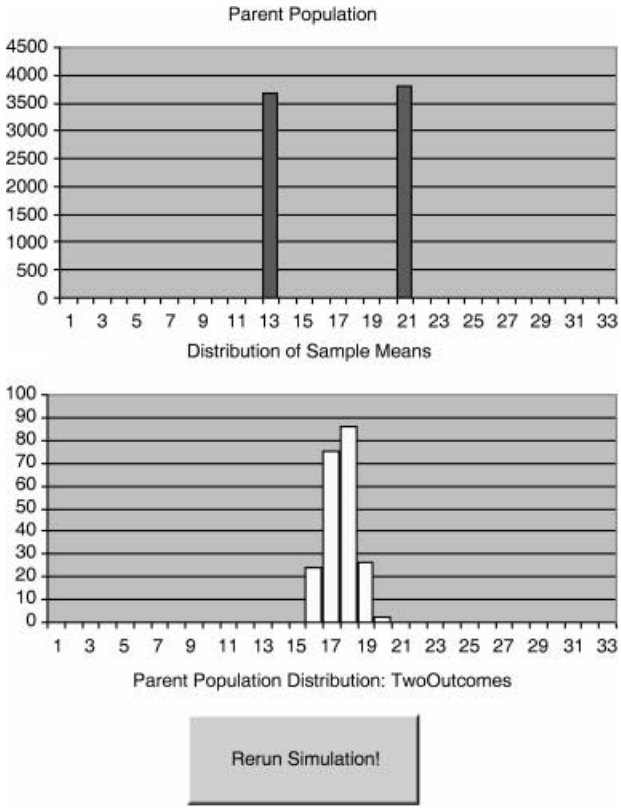

Fig. 4. A larger number of samples (250)

small sample size, gradually increasing it to build intuition. Finally, students were advised to change distributions only after fully exploring the effects of the other parameters using each one. We will discuss the students' written comments before turning to their performance on the statistical problems and quantitative evaluative questions.

As with our previous efforts at facilitating visual learning (see Jones et al. 2004), most students prefer the simulation to didactic instruction.

> 'I found it to be self-explanatory and helpful to work with because I am a visual learner and it is a good visual tool.'

Likewise, this software was often preferred to other software used in the class.

> '[Visual Sampling Distributions] is more concise and more focused towards what we do in class. It is also more universally usable to be able to run the simulations quickly and derive a result.'
>
> 'It is very helpful to be able to configure different samples and see the different results.'

However, it is important to emphasize that the software is most successful when introduced during classroom time and when students are given the opportunity to experiment with instructor supervision. Furthermore, some enterprising students were able to strain the numerical limits of Excel, and some Excel security issues presented challenges for technical support. While it may be valuable to develop the simulation as a stand-alone tool, in our view the ease of development and the wide accessibility offered by VBA outweighs the disadvantages.

In addition to the qualitative feedback, we also gathered usability and outcome data, including how much time students spent on the explanation (explain time), how much time students spent on experimenting with parameters (experiment time), how well the students felt the software explained the concept of sampling distributions (explain topic), how much they learned (learned experimenting), the value of the time required (time required) to use the program compared to their increased understanding, and how user-friendly (user friendly) the software was perceived to be. The last four questions were evaluated on a Likert scale of 1 to 7. We also asked students how long they had spent evaluating the software (evaluation time) and how much time they had spent on the substantive statistics questions (test time).

The resulting correlation matrix yields some interesting insights (see table 1).

1. Those students who invested the time to digest the instructions and to learn to use the program

© 2007 The Authors
Journal compilation © 2007 Teaching Statistics Trust

	Explain time	*Experiment time*	*Explain topic*	*Learned experimenting*	*Time required*	*User friendly*	*Evaluation time*	*Test time*
Explain time	1							
Experiment time	0.62	1						
Explain topic	−0.2	0.05	1					
Learned experimenting	−0.2	0.01	0.46	1				
Time required	−0.2	−0.1	0.47	0.66	1			
User friendly	0.05	0.13	0.4	0.14	0.21	1		
Evaluation time	0.42	0.12	−0.1	0.02	−0	−0.29	1	
Test time	−0	−0.2	−0.4	−0.6	−0.6	−0.23	0.26	1

Table 1. Correlations for the sampling program survey, evaluation questions

in fact spent time using the program, and learned by experimenting.

2. Also not surprisingly, if the student felt that the topic was not well explained or felt that he did not learn by experimenting, he did not benefit from the software and the statistical test took more time.
3. Time spent reading the explanation is not correlated with how well the topic was explained, or how much was learned by experimenting.

Finally, in order to determine whether or not there were learning gains that could be attributed to the new software, the students were asked to respond to questions designed to test substantive knowledge of relevant statistical concepts. See the Appendix for some examples of these questions. In fact, the students who evaluated this program did statistically significantly better than the control group on problems that asked about repeated sampling.

◆ CONCLUSION ◆

Many students, particularly those with a penchant for visual learning, appreciated having access to the new software, and it did appear to impact the students' understanding of the concepts of sampling distributions and the Central Limit Theorem. As any good teacher understands, learning is often accomplished by doing rather than by reading, or by mere listening. However, our analysis strongly indicates that these benefits will accrue only when classroom time is employed to convince the students of the efficacy of the simulation approach. But the investment is well worth it. Visual Sampling Distributions is another example of the power of resampling to illuminate difficult statistical concepts.

Because Visual Sampling Distributions is implemented in Visual Basic for Applications, it is easily modified to suit particular needs. Download of Visual Sampling Distributions and Visual Regression is available free of charge at http://law.gsu.edu/cncr/research/jones.html.

Reference

Jones, G.T., Hagtvedt, R. and Jones K. (2004). A VBA-based simulation for teaching simple linear regression. *Teaching Statistics*, **26**(2), 36–41.

Appendix

Sample substantive questions

A random sample of 40 observations is selected from a very large population of ball bearings, with diameters that are Normally distributed with mean 5.0 mm and standard deviation 0.001 mm. $\bar{X}$ denotes the sample mean diameter for the 40 ball bearings in the sample. Which of the following statements is true?

a. Because the sample is not large, we do not know the distribution of $\bar{X}$.
b. $\bar{X}$ is Normally distributed with mean 5.0 and standard deviation smaller than 0.001 mm.
c. $\bar{X}$ has a binomial distribution with $n = 40$ and $p = 0.001$.
d. $\bar{X}$ has a Normal distribution with mean 5.0 and standard deviation 0.001 mm.

If the sample size in the question above had been 50 ball bearings, the standard deviation of the $\bar{X}$ distribution would be

a. larger than the standard deviation of the $\bar{X}$ distribution where the sample size is 40.
b. smaller than the standard deviation of the $\bar{X}$ distribution where the sample size is 40.
c. could be larger or smaller than the standard deviation of the $\bar{X}$ distribution where the sample size is 40.

© 2007 The Authors
Journal compilation © 2007 Teaching Statistics Trust

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	