DETERMINING THE EDUCATIONAL POTENTIAL OF COMPUTER BASED STRATEGIES FOR DEVELOPING AN UNDERSTANDING OF SAMPLING DISTRIBUTIONS

PETER JONES & KAY LIPSON
Swinburne University of Technology

An analysis of the steps involved in forming the idea of an empirical sampling distribution and the nature of the methods and/or images used in most computer based strategies to teach this idea suggest that this way of using the computer adds little insight to the usual text based explanations that they are designed to complement. This analysis suggests reasons why a more recent approach which uses the computer to model and dynamically display the processes that underlie the idea is more likely to be successful.

A critical step in developing the theory of statistical inference is the idea of a sampling distribution – the recognition that the estimates of a population parameter will vary and that this variation will conform to a predictable pattern. Yet, for all its importance, experience and research have shown that the idea is generally poorly understood (Moore, 1992; Rubin, Bruce, & Tenney, 1990 for example). One reason for this might be the way in which the idea has been traditionally introduced in statistics courses using a deductive approach based on probability theory (Johnson & Bhattacharyya, 1987; Mendenhall, Wackerly, & Scheaffer, 1990 for example). Such explanations are usually expressed in a highly mathematical language which tends to make the argument inaccessible to all but the mathematically able, now a very small minority of the students taking introductory courses in inferential statistics. But perhaps more importantly, it is a theoretical development which is difficult to relate to the physical process of drawing a sample from a population. Statistics educators have come to recognise that there are deficiencies with a purely theory based explanation and now often accompany or replace this with an empirical argument. The alternative interpretation uses the long run relative frequency approach, where the sampling distribution is viewed as the result of taking repeated samples of a fixed size from a population and calculating the value of the sample statistic for each (Devore & Peck, 1986; Ott & Mendenhall, 1990 for example). The empirical approach has the advantages of being more readily related to the actual physical process of sampling and requiring minimal use of formal mathematical language.

Because the computer has an obvious role in the empirical development of the idea of a sampling distribution, by carrying out the repeated sampling and summarising the results, a number of instructional sequences have been developed built around these capabilities. Unfortunately these approaches, although widely promoted and now commonplace activities in introductory statistics courses, have been less successful than statistics educators might have hoped for, as noted by Hawkins (1990):

> *ICOTS 2 delegates were treated to "101 ways of prettying up the Central Limit Theorem on screen", but if the students are not helped to see the purpose of the CLT, and if the software does not take them beyond what is still, for them, an abstract representation, then the software fails. (p 28)*

In this paper we will look at why early attempts at computer based explanations have not been as successful as they might be, and suggest reasons why a more recent approach by Rubin and her colleagues (Rubin, 1991) is more likely to succeed.

THREE EMPIRICALLY BASED STRATEGIES FOR INTRODUCING THE IDEA OF A SAMPLING DISTRIBUTION

In this analysis we will consider three empirically based instructional strategies for introducing the idea of a sampling distribution. For convenience we will restrict ourselves to the distribution of a sample proportion. The first strategy is a typical text based explanation. The second strategy utilises the general purpose computer package Minitab. The third strategy involves a computer package which explicitly makes use of the increased graphics potential of the new desktop computers.

Strategy 1 (Text only)

Text based instructional strategies verbally describe the process of forming an empirical sampling distribution. These explanations are often accompanied by one or more relative frequency histograms showing the distribution of the sample proportion $\hat{p}$ for a large number of trials, typically several hundred. The reader is then asked to note the (generally) near normal shape of the distribution and that it is centred on or around the (known) population proportion p. Figure 1 is taken from such a development sequence in a typical introductory statistic text (Devore & Peck, 1986 p 255). The population under consideration here is the labour force in Ireland, and the population proportion p = 0.265 is the percentage of females in this population.

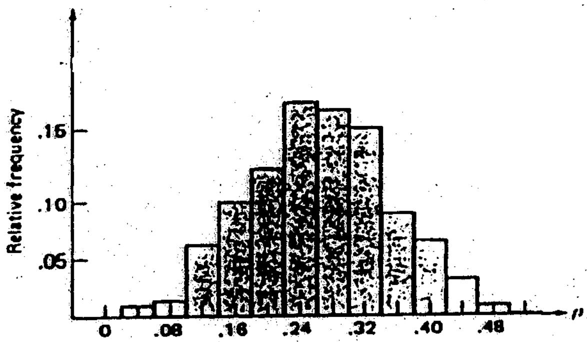

Figure 1: Histogram of 500 values of $\hat{p}$ based on a random sample of size n =25 (p=0.265)
(Devore & Peck, 1986 p 255)

Strategy 2 (Minitab)

In the early 1980's the more innovative statistics educators began using the computer as part of their teaching sequence (Bloom, Comber, & Cross, 1986; Thomas, 1984 for example). In the earliest attempts complicated programming was required, but now commonly available statistical computer packages such as Minitab may be used to produce empirical sampling distributions. Students are given the appropriate computer code to generate random samples, calculate corresponding values of the sample proportion $\hat{p}$, and display the distribution graphically (generally in the form of a histogram). For example, using Minitab, a histogram similar to that shown in Figure 1 can be reproduced using the following commands:

```
MTB > random 100 c1-c25;
SUBC> bernoulli 0.265.
MTB > rsum c1-c25 c30
MTB > let c40=c30/25
MTB > ghist c40
```

Using similar commands we can create any other histogramswe might like by varying the population proportion p, the sample size n, or both.

The resulting histogram is shown in Figure 2.

Strategy 3 (Sampling Laboratory)

More recent computer applications in mathematics and statistics have tended to de-emphasise the use of the computer as a computational tool and focus on our additional ability to use current technology to build a working model of the process under consideration and to display the results graphically, for example

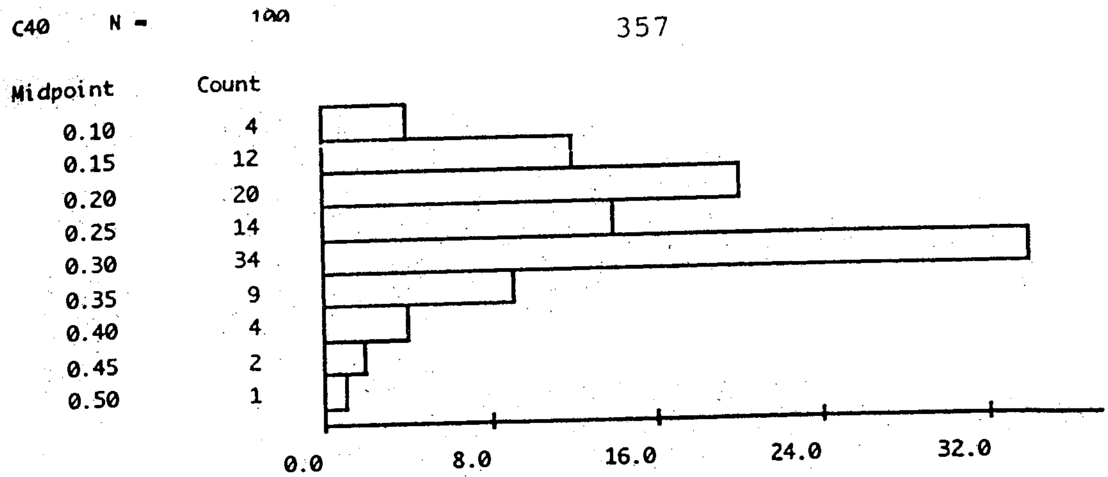

Figure 2: Histogram of sampling distribution produced by Minitab

Geometric Supposer (Shwartz & Yerushalmy, 1985). Such an approach has been followed in Sampling Laboratory by Rubin et al. To use the package no programming is required by the user who simply makes the appropriate entries as requested. To replicate the example previously used the user enters the name the population (Labour Force Ireland), the attribute of interest (female, male) and the appropriate value of p (0.265). 'Experiments' are then invited, with the student requested to name the experiment (n=25), enter the sample size (25) and the number of samples to be drawn (100). The required number of samples are drawn sequentially, and the screen shows simultaneously the following three windows:

1 A probability distribution/bar chart of the population proportions.
2 A histogram which shows individual sample outcomes as well as the number of each.
3 An empirical sampling distribution of the values of the sample proportion $\hat{p}$ which builds as the sampling proceeds, with the value of the sample proportion $\hat{p}$ from the last sample shown explicitly in black, and also the overall sample proportion.

The 'Experiments' screen after 29 samples have been drawn is shown in Figure 3.

The same calculations which are performed in Minitab are also carried out here, but the emphasis is on the sampling process and the calculations remain very much in the background. In Sampling Laboratory the sampling process can be observed in real time and students see the sampling distribution form as more and more samples are taken. The process may be paused and restarted at any time, or may be conducted stepwise, one sample at a time.

Comparative educational gains through computerisation

In order to compare potential educational gains made by introducing the computer into the instructional strategies considered here it is useful to first identify the various steps involved in developing the idea. This has been done in the following table which also indicats the method or image primarily used to represent each step in the three instructional strategies. Note that we would not expect the computer based strategies to stand alone, but to be considered as complementary to the text based strategy. Thus, any representations available using the text based strategy would also be available when using the computer based strategies.

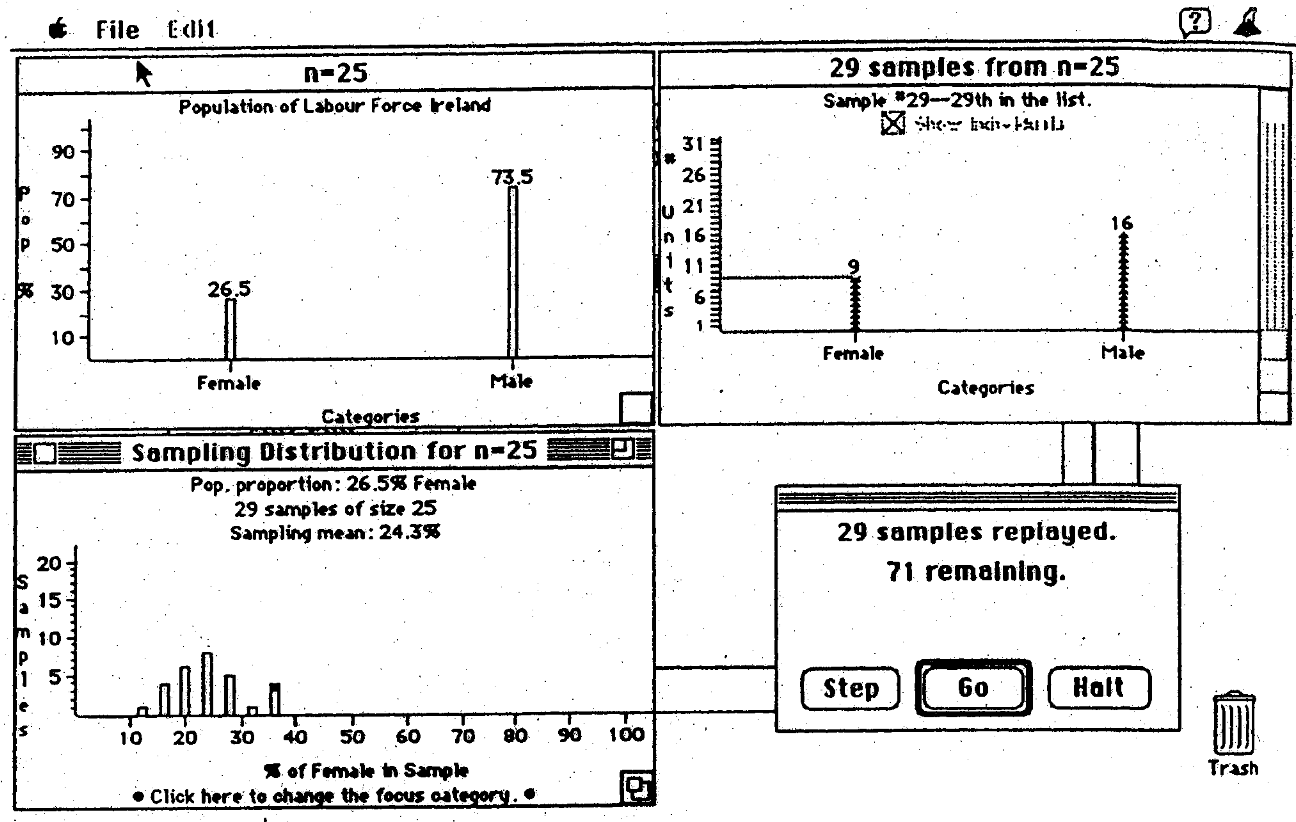

Figure 3: The 'Experiments' screen from Sampling Laboratory after 29 samples have been drawn showing the population proportion, the last sample and the empirical sampling distribution.

Development Sequence	Strategy 1 (Text)	Strategy 2 (Minitab)	Strategy 3 (Samp Lab)
1 A sample of individuals of given size is chosen from a *population*.	written explanation	implicit in code	population named, sample size specified
2 Those chosen in the *sample* may exhibit many attributes, one of which is to be recorded.	written explanation	implicit in code	attribute name specified
3 The proportion of individuals in the population which possess the attribute of interest is constant.	written explanation	implicit in code	the population proportion p is specified
4 The number of individuals in the sample which possess this attribute can be calculated.	written explanation	implicit in code	sample outcomes are depicted in a histogram
5 Many different samples can be selected from a population.	written explanation	samples can be listed and compared if requested	dynamic display of sample outcomes show variation from sample to sample

6 Each sample gives rise to a value of the sample proportion $\hat{p}$.	written explanation	the value of $\hat{p}$ can be printed if requested	the value of $\hat{p}$ is shown on the histogram
7 These values of the sample proportion $\hat{p}$ will not all be the same.	written explanation	all values of $\hat{p}$ can be listed if requested	dynamic histogram which accumulates all values of $\hat{p}$
8 The values of the sample proportion $\hat{p}$ will form a distribution, called the (empirical) sampling distribution of $\hat{p}$.	written explanation and static histogram	static histogram	dynamically formed histogram of values of $\hat{p}$

Discussion and Conclusion

What then has the introduction of the computer contributed to the development of the idea of a sampling distribution? An analysis of the steps involved in developing the idea of a sampling distribution shows that it is a complex idea whose understanding cannot be dissociated from an understanding of the sampling process by which it is formed. The sampling process is dynamic and involves the linking of several elements: a parent population, the samples drawn from the population, the values of the test statistic extracted from each of the samples, and the sampling distribution they give rise to. In a purely text based explanation, the written word is used to describe the process outlined above and graphics are used to illustrate the end product of the process, the empirical distribution of the sample statistic. The bringing together of the sampling process and the resulting sampling distribution requires a high degree of mental processing and visualisation which seems to be difficult for many students because of the high cognitive load it potentially imposes.

However, an analysis of the second strategy shows that the introduction of the computer has done little to ameliorate the problem. Certainly, the computer gives the student an active role to play in generating the samples, extracting the values of the test statistic and forming the sampling distribution. It also gives them the opportunity to experiment with parameter values of their own to see how the form of the sampling distribution changes in response. But what has been done by the computer could have been achieved, in theory at least, with pen and paper. The product (the empirical sampling distribution) is certainly there but the process involved in generating the sampling distribution is hidden in the code (see previous section). As a result, it would appear that this computer based strategy would do little to reduce the high cognitive load that is inherent in a purely text based explanation. This contention would appear to be supported by classroom experience.

The third strategy utilising the computer package Sampling Laboratory creates a working model of the sampling process which has as its product a histogram displaying the resulting sampling distribution (similar to that produced by Minitab), but also displays the process by which the sampling distribution is obtained. This is done by simultaneously displaying the parent population, the current sample, the state of the sampling distribution with the addition of information from that sample, and finally the way in which each of the elements changes as the sampling process is set in motion. While graphical images similar to those that make up this display could have been generated by Minitab on an individual basis, the power of computer based representation of Sampling Laboratory is the simultaneous display and interlinking of these images so that changes in one element are immediately reflected, where appropriate, in the other elements. These dynamic, interactive displays have the potential to considerably reduce the cognitive load involved in understanding text based explanations of the idea of a sampling distribution and should lead to greater understanding. However, this has yet to be empirically tested.

The analysis so far has been concerned with a particular instructional problem, the understanding of the idea of an empirical sampling and specific technology, Minitab and Sampling Laboratory. However, as indicated by Kaput (1992), if we are to make progress in understanding the way in which technology can be integrated into the educative process and not just replicate what we have done in the past, we need to be able to step back from specific content areas and technologies and look at the general principles involved. Two of Kaput's key principles emerge from this analysis. Firstly, with computers we are now in the position to create new notations (ways of recording and /or displaying information) (Kaput, 1992 p523) that are more capable of conveying a complex idea than the traditional paper based notations used currently. This has been clearly illustrated in Rubin's Sampling Laboratory. While packages such as Minitab could actually generate similar graphical images to those used by Sampling Laboratory, even if the individual images produced could be arranged appropriately and displayed simultaneously, the end result would be no different than printing out the same images on a sheet of paper. To use Kaput's terminology, the notations used by Minitab, like those of most other statistical packages, have not been designed to do any more than replicate paper based notations which in turn have been designed to convey information in an inert medium. Whilst there may be an educational payoff of using the computer in this instance, by enabling the student to be more actively involved in the learning process, the images produced convey no additional information to the student.

The second general point made by Kaput and others (Lesh, Post, & Behr, 1987 for example) is the need to recognise that many concepts in quantitative disciplines like mathematics and statistics are multi-faceted and cannot be fully understood from a single representation. The sampling distribution is such a concept. Its definition as the distribution of a combination of random variables is not easily related to the physical process of sampling. Similarly, the empirical approach, which sees the sampling distribution as a long term frequency distribution, captures this aspect but does not lead to the precise relationships specifying the centre and spread of the sampling distribution that are the corner stones of almost all introductory statistical inference courses. Kaput suggests that to fully understand a concept it is necessary to understand it in all its manifestations. In terms of the sampling distribution, this would call into question any instructional sequence that presented only one approach, in this case purely theoretical or purely empirical. In this regard, neither of the computer based instructional sequences presented here come to terms with the issue.

REFERENCES

Bloom, L. M., Comber, G. A., & Cross, J. M. (1986). Using the microcomputer to simulate the binomial distribution and to illustrate the central limit theorem. International Journal of Mathematics Education in Science and Technology, 17(2), 229-237.

Devore, J., & Peck, R. (1986). Statistics. St Paul: West Publishing Company.

Hawkins, A. (1990). Success and Failure in Statistical Education - A UK Perspective. In D. Vere-Jones (Ed.), Proceedings of the Third International Conference on Teaching Statistics, 1 (pp. 24-32). Dunedin, NZ: ISI.

Johnson, R., & Bhattacharyya, G. (1987). Statistics Principles and Methods. New York: John Wiley & Sons.

Kaput, J. J. (1992). Technology and Mathematics Education. In D. A. Grouws (Eds.), Handbook of Research on Mathematics Teaching and Learning (pp. 515-555). MacMillan.

Lesh, R., Post, T., & Behr, M. (1987). Representations and Translations among Representations in Mathematics Learning and Problem Solving. In C. Janvier (Eds.), Problems of Representation in the Teaching and Learning of Mathematics (pp. 33-40).

Mendenhall, W., Wackerly, D. D., & Scheaffer, R. L. (1990). Mathematical Statistics with Applications (Fourth ed.). Boston: PWS-Kent Publishing Company.

Moore, D. S. (1992). Teaching Statistics as a Respectable Subject. In F. Gordon & S. Gordon (Eds.), Statistics for the Twenty-First Century (pp. 14-25). MAA Notes No 26, Mathematical Association of America.

Ott, L., & Mendenhall, W. (1990). Understanding Statistics (Fifth ed.). Boston: PWS-KENT Publishing Company.

Rubin, A. (1991). Using Computers in Teaching Statistical Analysis: A Double Edged Sword. Unpublished Manuscript.

Rubin, A., Bruce, B., & Tenney, Y. (1990). Learning about sampling: Trouble at the core of statistics. In Proceedings of the Third International Conference on Teaching Statistics, 2 (pp. 314-319). Dunedin, NZ: ISI.

Shwartz, J., & Yerushalmy, M. (1985). The Geometric Supposers. Pleasantville, NY: Sunburst Communications.

Thomas, D. A. (1984). Understanding the central limit theorem. Mathematics Teacher(October), 542 - 543.