WIZARDRY OR PEDAGOGY?: WHAT IS THE DRIVING FORCE IN THE USE OF THE NEW TECHNOLOGY IN TEACHING STATISTICS

James Nicholson
Queen's University of Belfast
Gerry Mulhern
Queen's University of Belfast, Nth Ireland
Neville Hunt
Coventry University
United Kingdom

The availability of technology opens up opportunities for students to explore larger datasets and to gain experience of the effects of random variation. We have been involved in a development project to produce materials, with a sound pedagogical basis, to support the construction of accurate conceptual understanding of key statistical concepts. This paper presents the range of materials from the project and outline the pedagogical basis for them in light of the question posed in the title.

INTRODUCTION

The practice of statistics now is dominated by computer-based technology. A medium specification desktop computer now might typically have a 1.5 GHz processor with 256 MB of SDRAM, allowing computationally intensive methods to be undertaken personally rather than being referred to a specialist computing service. Technology has made equally large strides in other areas. For example, the quality and reliability of computer driven presentations have improved dramatically, and at the same time have become steadily more affordable and more manageable.

Computer-based learning (CBL) environments have also become much more accessible as a result of the hardware and software developments. At all levels and ages, students today will have had much more experience in the use of computers than their predecessors and, as a consequence, their keyboard and computer management skills are considerably higher.

Technical competence is required so that the use of the technology does not itself form a barrier to learning – the student can focus on the content of CBL or the outcomes of an analysis rather than having to concentrate on the mechanics of operating the computer. While significant improvements in students' technical competence have been achieved, there remain issues in respect of equality of opportunity in access to technology – the rate of progress for those already well-off is quicker than those with less provision so the gap is getting wider, both globally and within technologically advanced countries like the UK. Hawkins (1997) observed that 'Having the vision to see what technology can, or might, do is not synonymous with knowing how to take advantage of this in a teaching context.' The use of technology in teaching statistics therefore has a number of inter-related strands:

- issues regarding the use and usability of software packages.
- new opportunities to present ideas dynamically and interactively to students , rather than in more conventional, 'static' environments.
- the need to develop reliable models of conceptual learning and understanding to underpin the development of CBL.

ADVANTAGES AND DISADVANTAGES

While we would not advocate that users need to know the detailed workings of statistical algorithms employed, we would argue that they need to know something of the principles being used. For example, in working with lines of regression the user should understand that the observed data are approximated to a *straight* line using a specified criterion. It is important that statisticians understand the mathematics behind the derivation, but not that intelligent users of statistics do. Batanero, Godino, Vallecillos, Green, & Holmes. (1994) identified a wide range of conceptual difficulties experienced by students as statistics has become a much more common part of the standard curriculum internationally. Some of those difficulties are at the computational level, i.e. they relate to procedural competence, while others are epistemological, revealing a lack of understanding of the nature of the concept in question. Where calculation is automated, there are few is any procedural errors – for instance in calculating means or standard deviations at the secondary level, through to calculating *p*-values for more sophisticated tests or undertaking a factor analysis of multivariate data. Also much more is possible now in any given amount of time than previously.

This 'progress' may however be a double-edged sword: while technology enables a shorter time to be spent in doing the tedious calculations and manipulations which has characterised much of the taught statistics curriculum for a long time, the greater procedural competence may mask a lack of understanding of underlying concepts. Indeed, increased automation has inevitably reduced students' 'contact' with the actual data. Our experience in classrooms suggests that this lack of contact results in the data becoming rather anonymous, and abstract, to the student. As a consequence, a grasp of the meaning of the data and the sensibleness, or otherwise, of possible relationships is not well formed. It is tempting to assume everything is fine because students appear to be more accurate, but their understanding may be deteriorating because of this compression in time spent acquiring new schema – skipping Vygotsky's zone of proximal development, so to speak.

Moreover, the result of this technological progress is that sophisticated statistical analyses are no longer only accessible to those who are expert in both statistics and computer use. However, the software packages do not prevent inappropriate graphs being drawn, unjustified statistical procedures being applied to data, nor do they provide either the depth of inferential insight or the background knowledge [context specific specialist knowledge] needed to make reasonable interpretations of the analysis.

TYPES OF DIFFICULTY

What are the problem areas to which particular attention needs to be paid?

1) Poor graphical displays: while these have improved over time, commercial software still displays graphs with technical errors. To make the software more user-friendly, many packages will insert a scale automatically to maximise the spread of the display, irrespective of any distortion introduced in the perception of the data. Since the amount of work required to produce graphs is not an issue, it is tempting for users to go for the most colourful or most flamboyant. They may choose to display multiple graphs relating to the same information, rather than communicating the information most effectively.

	Change in pass rates
English	10%
Maths	-5%
Science	12%

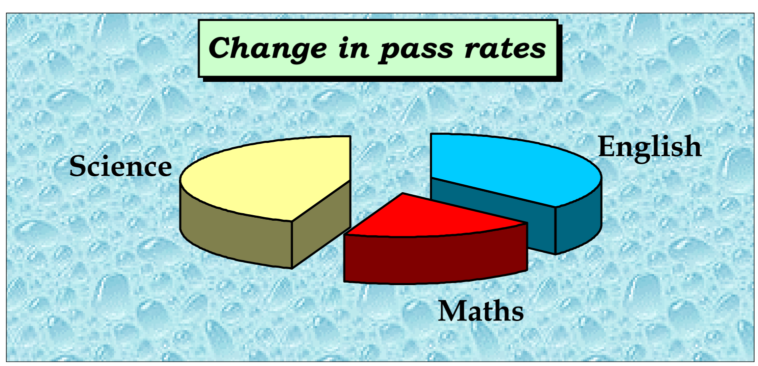

Figure 1: Pie Chart for % Data.

Students also fail to appreciate the subtlety of the meaning of concepts at times. For example, they learn that pie charts are most appropriate for showing percentages. If they then have a set of data where the values are percentages, they may choose a pie chart when it is not appropriate because they fail to grasp the distinction between data that represent percentages of different quantities, as in the case illustrated in Figure 1 above, and a graphical representation

showing each data item as the proportion of the total of those data items. Moreover, in this situation, a package such as Excel will happily ignore a negative sign indicating a decrease and draw an impressive looking pie chart which is meaningless.
2) Inappropriate analysis: menu-driven programs are constructed so that certain options are not available (usually appear as grey rather than black) if they cannot be applied at that point. This is straightforward in many contexts – in word-processing, ‘cut’ and ‘copy’ only become active when some portion of the document has been selected. However, it is less straightforward in many statistics contexts. Data given as %’s should not be subjected to a Chi-squared test, since the size of the sample is of critical importance, but the computer will see numerical values in the table, and offer a chi-square analysis as though they represented counts, as in Figure 2 below. Note that Minitab’s output actually refers explicitly to ‘counts’, but in a form, which is unlikely to alert the unwary to a potential problem.

Chi-Square Test: Con, Lab, Lib

```
Expected counts are printed below observed counts

           Con      Lab      Lib    Total
    1       43       62       46      151
         50.33    50.33    50.33

    2       40       30       30      100
         33.33    33.33    33.33

    3       17        8       24       49
         16.33    16.33    16.33

Total      100      100      100      300

Chi-Sq =  1.068 +  2.704 +  0.373 + 1.333 +  0.333 +  0.333 +
          0.027 +  4.252 +  3.599 = 14.023
DF = 4, P-Value = 0.007
```

Figure 2: Minitab Output for an Inappropriate Chi-Square Analysis.

Or the analysis may simply be meaningless. Data can now be collected very cheaply, and students, and professional users of data, may be faced with a large data set with many variables. The data may well show a ‘significant difference between height and IQ’ for example, if a test comparing means is carried out. Furthermore, because of the structure of hypothesis testing, you can generate a proportion of ‘significant’ results by undertaking a sufficient number of statistical tests on data even if the population does not possess any of the characteristics of interest. This tendency towards ‘statistical napalm’ is perhaps the most common problem arising from access to powerful, IT-based statistical packages.

At a more subtle level, the data analysis may have meaning, but there may be confounding variables which the computer won’t discern – i.e. Simpson’s Paradox. This latter difficulty is by no means unique to the use of technology, but the facility with which the analysis may be done by computer may encourage less focus on such issues. We would want to encourage students to use the reduction in time required to learn complex manipulative procedures to spend more time thinking about the data, and the contextual issues, but this is not easy to accomplish, and there is a real need for us to understand better how to help develop sophisticated reasoning skills.

ILLUSTRATION OF PEDAGOGICAL POSSIBILITIES

It is likely that everyone would agree that pedagogy *should be* the answer to the question posed in this article’s title, but is it? If we just use PowerPoint to present lessons or lectures more attractively, or to produce copious quantities of output viewing the data from every conceivable angle, are we exploiting the potential of the new technology fully or appropriately?

The combination of new presentation hardware and software with software packages that can record the history of an analysis opens up new possibilities. There is more than one level at which we might want to engage a student’s attention with this output – for example in Time Series analysis, typically, a series of possible models may be examined to see how each fits the data. Initially a teacher might highlight the procedural aspects: the software performs a number of tests on the model, and provides a number of summary statistics. The teacher might draw attention to the detail of where the p-values have come from, so that students are aware of the information contained in the output, and how it is constructed. However, there are a number of criteria used in evaluating the models in time series, and in most instances there is not one model that provides the best fit for all criteria. The same output may then be viewed at a different level, and attention drawn to the elements in the output which will be used in evaluating one model compared with other possible models to illustrate the appropriate decision making processes.

Simulations and visual representations of statistical concepts are not new opportunities in statistical education. They have been around for some considerable time. Seeing the shape of a Binomial or a Poisson distribution, and how they depend on the parameter values is a good example of a simple, yet very powerful, visual representation. Good teaching harnesses such tools to develop sound understanding of more complex ideas, such as approximating the Binomial by the Poisson. The general notion of approximating distributions is not trivial, since the nature of the parameters (their physical meanings in context) are usually quite different, and direct association is therefore difficult. Being able to manipulate the parameters of two or more distributions displayed simultaneously can allow the user not only to develop an understanding of the parameter values (in the approximating distribution) which give the best possible fit, but also an appreciation of how good the approximation is – and that there is actually a continuum from very poor fit to excellent fit in which we, rather arbitrarily, choose a cut-off.

Greer (2000) observes 'As a result of technological developments, the ratio access to data analytical and critical tools for interpretation is accelerating out of control' (p.XX). Since the 'access to data' part of this ratio is likely to continue to increase exponentially, it is crucial that we improve the interpretative tools and skills faster as well. The materials we have been developing try to provide a sound grasp of the nature of variation which is a core conceptual obstacle to interpretation and inference.

'DISCUSS' – SAMPLING AND ESTIMATION

The role of variation is perhaps the most fundamental component in understanding and interpreting data. Most textbooks deal with quantifying it with measurements such as standard deviation, range, inter-quartile range … and there is little discussion about the sources and causes of variation. Variation exists in all measurable quantities. Sources of variation are complex and contextually dependent, making it hard for the student to achieve a sound construction of the concept. For example, in a manufacturing process the quality of raw materials, the skill of workers and the level of monitoring during production may be the key factors, while in plant growth, the physical environment, the soil quality, the weather, genetic make-up and the time of day of measurement may play a comparable role.

The difficult task in analysing data is distinguishing where observations are consistent with the natural variability in that context, and where they signify that something more substantial has changed. Our 'Sampling and Estimation' materials seek to develop an experience based understanding of the amount of variation that may be expected in different contexts. We have tried to make them as accessible as possible in the current technological environment, by using a web-based structure. Hunt and Tyrrell (2000) provide a fuller discussion of the technical aspects of this material. While the simulations run in Excel, the users do not have to know Excel in order to work with them. Users run the simulations, controlling various parameters to explore how the behaviour changes, while being guided through the activities by an on-screen worksheet. Linkages of web pages provide contextual support and explanation.

We have aimed to make navigation within the site as straightforward as possible, so that users can easily access any particular section if required, but also to provide a clearly visible 'recommended route' which leads new users through the development of new ideas in a manner which we judge to be coherent. The first level explores the behavior within samples, using different population distributions to draw out similarities and differences that exist in that behavior. Sample statistics are shown in later simulations so as to highlight some of these features more fully. Figure 3 shows an example of the simulation screen for a sample from the exponential distribution showing the position of the mean. The accompanying activities ask the user to take a number of samples, and consider the position of values within those samples, to build up an understanding of behavior within a skewed distribution, and compare and contrast this with the behavior when sampling from the uniform or normal distributions.

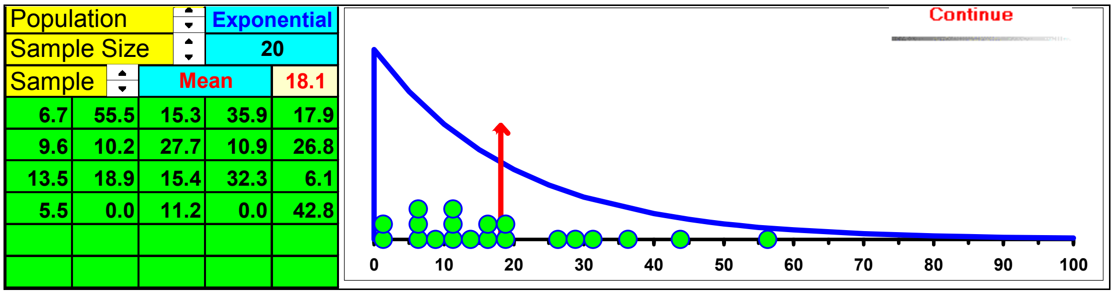

Figure 3: Values Sampled from an Exponential Distribution.

Users progress to exploring the behavior of the statistics derived from various distributions and build up an experience-based feel for the variation which is to be expected. The outcomes of some of these simulations can then be used to explore the features of both the distributions and the estimators, which give rise to certain striking results. Figure 4 below shows the sampling distributions for mean, median and mid-range statistics for a uniform population. Students are used to the mean being the statistic of choice, yet here the mid-range is patently a much better estimator. The most effective use of these materials is when the users engage with them individually or in pairs, followed by a discussion in a larger group, moderated by a teacher. In such a setting, the reason why is accessible to any moderately able student, and draws out a fuller understanding of both the distribution and of what is happening in using these three statistics.

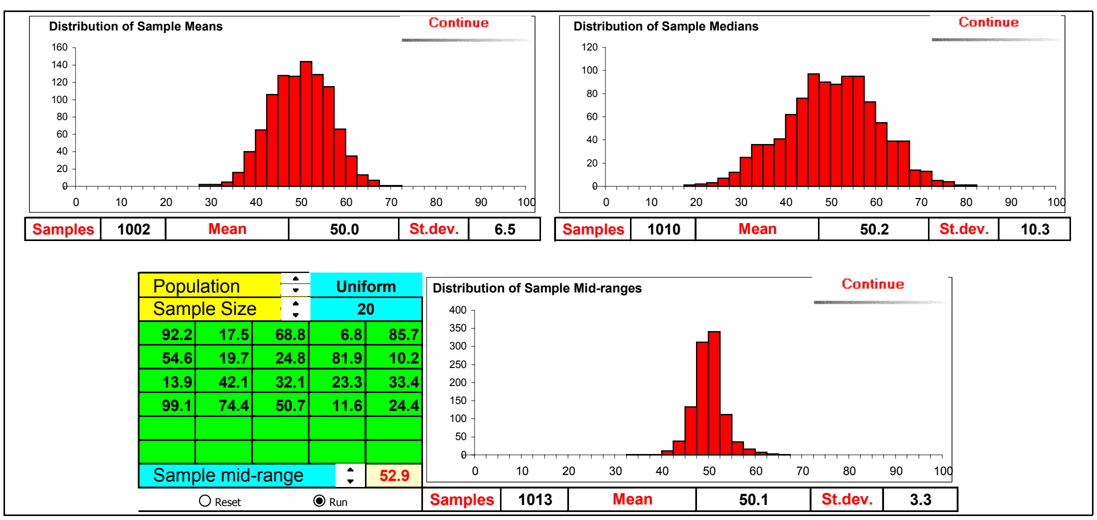

Figure 4: Estimator Behaviour with the Uniform Distribution.

If the same comparison is undertaken with a skewed distribution, other concepts such as the need for an estimator to be unbiased can be drawn out, and the effect of outliers on different statistics becomes clearer. Figure 5 below shows the same sampling distribution for an exponential population.

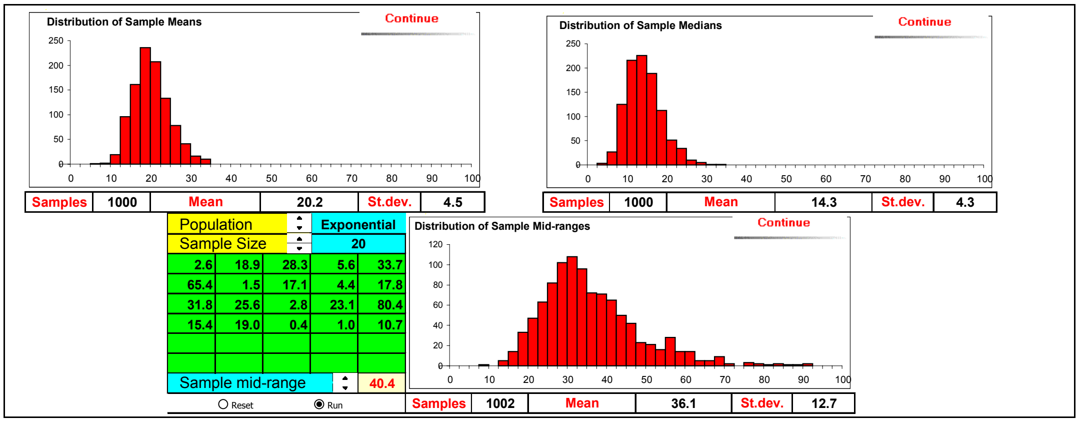

Figure 5: Estimator Behaviour with the Exponential Distribution.

Copies of the materials are available free by email request from the first author.

CONCLUSIONS

The use of dynamic, rather than static, presentation of images represents a paradigm shift, and opens up new opportunities for the instructor to bring ideas to life. Is it always a bonus? The danger exists that students do not grasp the connection between the representations being portrayed dynamically, perhaps the link between parameter values and the outcomes which are dependent on them. Even where these links are deterministic or functional, students may fail to appreciate the roles of the parameters. Where the outcomes are subject to random variation, the connections require much greater conceptual insight and if the instructor does not mediate effectively then the experience may be counter-productive in the sense of introducing greater confusion.

Enthusiastic disciples advocate the widespread adoption of the new technology. However, we need to recognise that new strategies are required. The majority of teachers of statistics at school level come from a mathematics background, and very many statistics courses at college level are taught within other academic disciplines by those with little formal training in statistics, and without much specialised training in the use of technology. We believe there are opportunities for CBL to contribute to statistical understanding, and there is a need for the development of more materials with sound pedagogical underpinnings to exploit those opportunities.

ACKNOWLEDGEMENT:

This work was partially supported by a grant from the Nuffield Foundation.

REFERENCES

Batanero, C., Godino, J.D., Vallecillos, A., Green, D., & Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics, Education, Science and Technology, 25, 527 – 547.*

Greer, B. (2000). Statistical thinking and learning. *Mathematical Thinking and Learning, 2, 1 – 9*.

Hawkins, A. (1997). Myth-Conceptions! In J.B.Garfield & G. Burrill (Eds). *Research on the Role of Technology in Teaching and Learning Statistics.* Voorburg: International Statistical Institute.

Hunt, N., & Tyrrell, S. (2000). Learning statistics on the web – DISCUSS. *Teaching Statistics*, 22, 85-90.