3. DEVELOPING PROBABILISTIC AND STATISTICAL REASONING AT THE SECONDARY LEVEL THROUGH THE USE OF DATA AND TECHNOLOGY

James Nicholson
Belfast Royal Academy

INTRODUCTION

Technology offers an end to the tedious and laborious computations in data analysis, but it also offers the possibility of a total lack of feeling for what is being done in the analysis, and a blind assumption that if the computer or calculator has done it then it must be right.

However, we can make use of technology, and realistic datasets, to enlarge our students' horizons in various ways.

- We can provide students the experience of **how random** random events are; for example, by having students analyze large datasets, conduct simulations, and generate samples from large distributions. We can use realistic datasets with students throughout the secondary level to develop their critical evaluative skills over a period of time.

- We can give students experience on which to build their intuition about what is going on in some of the sophisticated and some of the not so sophisticated analyses that the technology will do for them. We can illuminate some difficult concepts, including some in which the initial effect of using technology is often to produce misconceptions.

I work with students ages 11 - 18 in a selective (grammar) school of about 1,400 students. In November, 1995, my classroom was equipped with a computer and LCD panel, which has enabled me to use a dynamic blackboard during lessons and in the computer lab. This has been particularly useful because of the constraints on getting to a computer lab, which have been very busy with scheduled *Information Technology* and *Computer Studies* classes. As of September, 1996, I hope to have better access to a lab, because another lab is being created that will not have regular classes scheduled in it, but will be available for booking on a week-to-week basis. However, the *Information Technology* class is identified as a cross-curricular theme in the Northern Ireland curriculum; thus, all subjects have *Information Technology* classes and so the demand on a computer laboratory could be very high in a school the size of the Academy.

CRITICAL EVALUATION

Statistical packages and straightforward spreadsheets now offer very powerful charting facilities that take the hard work out of the presentation of information in graphical form. In EXCEL, which produced the charts used in the illustrations below, the *Chart Wizard* facility allows a chart to be drawn with virtually no effort--highlight the section of data to be presented; click and drag to position and size the chart; and then answer a series of questions as to the type of chart, whether the data is in rows or columns, and how the data, axes, and chart is to be labeled. By doing this, EXCEL produces a wonderfully professional looking graph; however, EXCEL does exactly what you tell it, and cannot tell if the data makes sense, or if the type of graph chosen is appropriate. Somehow we need to develop in the students, at an early age, a critical faculty to evaluate different forms of presentation, so that they can reject those that are definitely inappropriate and select the most informative graph when there is more than one appropriate possibility.

The charts shown here all relate to the simple dataset shown in Table1, which reports the sources of revenue for a Rugby Club over the first quarter of the year. Figure 1 shows this in the form of a three-dimensional comparative bar chart. Figure 2 is a two-dimensional stacked bar chart. Figures 3-6 report the four sources of revenue as pie charts. One problem becomes immediately apparent when looking at the four pie charts together--all of them have one sector representing 50% of the dataset. One would like to think that anyone drawing these would be struck by the visual pattern and question why this is so. Note that it is a trivial matter to go back into the chart wizard to use only the first three columns of figures. Even with a set of four pie charts presented as in Figures 3-6, many students do not query its appropriateness. If only one pie chart is used, then very few students feel that there is something wrong. The stacked bar chart in Figure 2 has already made this adjustment--only the first three columns are needed because the fourth column is represented by the overall height of the bar in each case.

Table 1: Sources of revenue for the rugby club

	Jan	***Feb***	***March***	***Total***
Advertising	750	500	800	2050
Sponsorship	1150	1200	1200	3550
Ticket Revenue	2450	2050	3200	7700
Totals	4350	3750	5200	13300

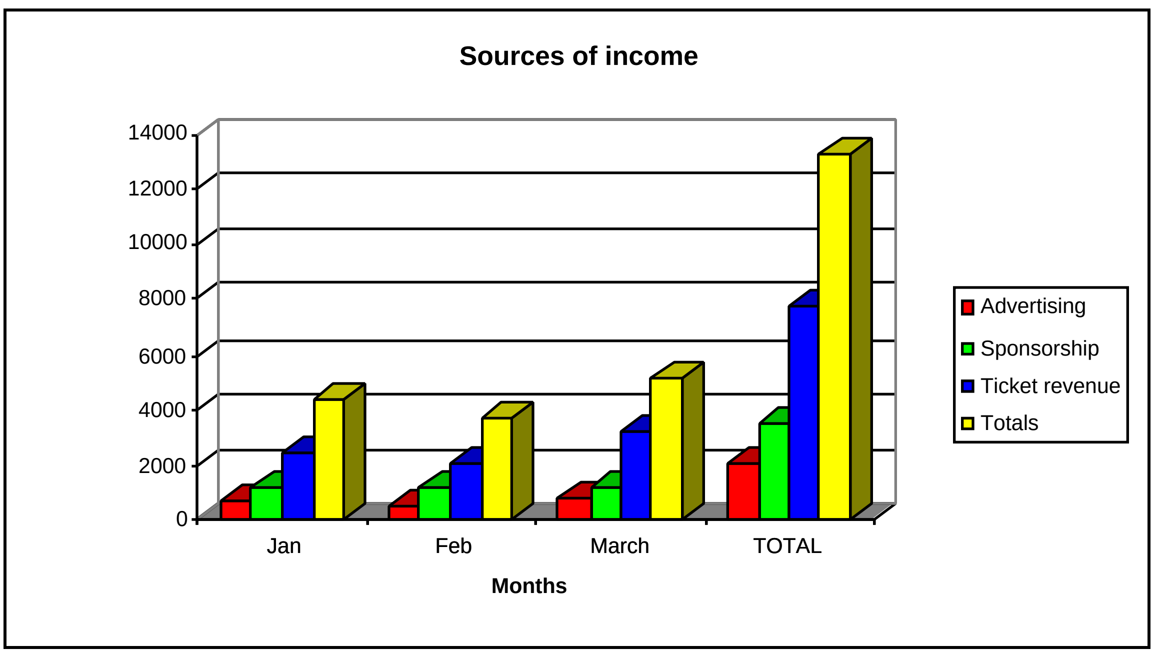

Figure 1: A three-dimensional comparative bar chart

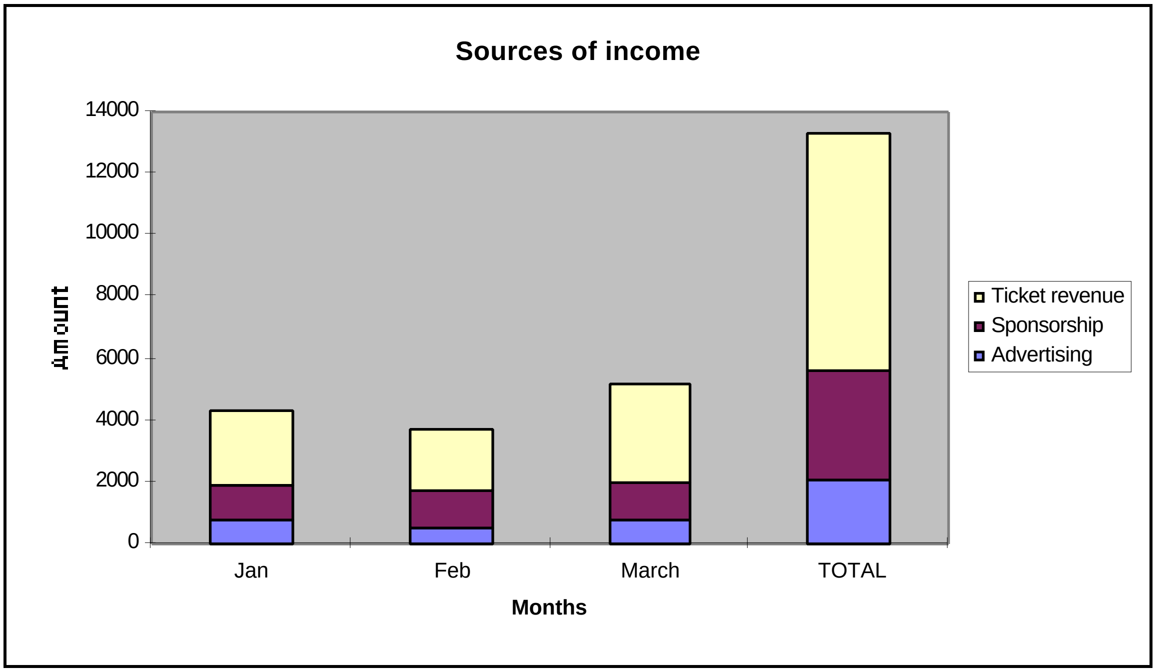

Figure 2: A two-dimensional stacked bar chart

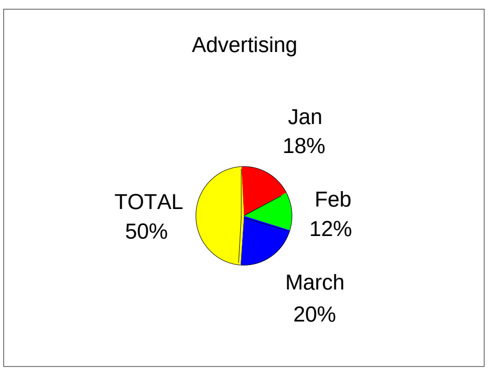

Figure 3: Advertising revenue

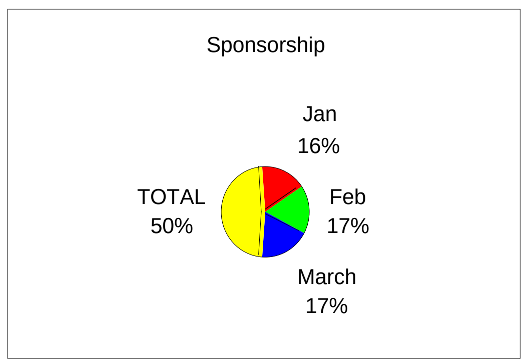

Figure 4: Sponsorship revenue

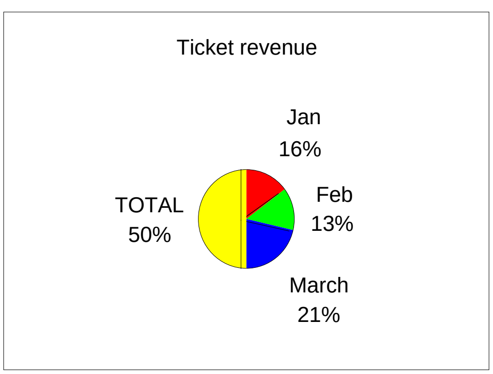

Figure 5: Ticket revenue

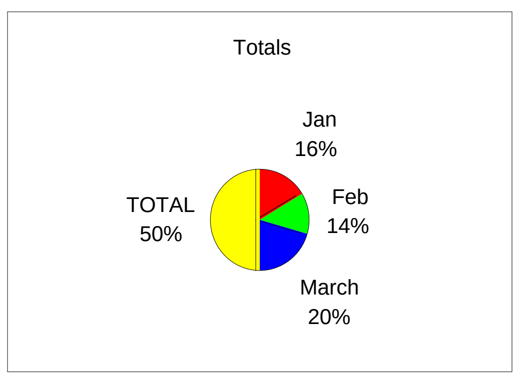

Figure 6: Total revenue

The experience of using these tools can help develop critical faculties, provided that the environment for exploring with them is reasonably focused, particularly when a student is just beginning to use them. There is no need for the datasets to be complicated; indeed, I think there is a great deal to be said for using extremely simple datasets like this one when the most basic principles are being assimilated. The age at which the students are able to comprehend the dataset needs to be considered carefully, because if we present complex datasets that are beyond the students' comprehension, they have no way of evaluating the presentation of the chart.

In the past, I have found that a large number of students attempting to study statistics at a more advanced level have found it difficult to provide adequate interpretations of their results within the context of the problem situation. I would expect that students who have had experiences such as those described above, while they are still young, would become better communicators of formal statistical conclusions, but this is an area where I think there is a need for further research.

ORAL PRESENTATIONS

Many students seem to feel that interpreting datasets is very difficult. They look for subtle nuances and ignore the glaringly obvious, or they confuse correlation with causation. Encouraging students to present their conclusions to their peers, and using technology to instill confidence in the look of their presentation, whether it be oral or in a poster format, both help students gain an understanding of how to communicate in statistics.

In my third form (Grade 9) class, students had to conduct a statistical investigation. They were given some possibilities, but were encouraged to choose their own. The students' choices ranged from investigations of videos to examination marks, some used secondary data and others collected their own primary data. The group of students were in a high academic stream, by general ability, with quite varied motivation in mathematics. I chose the investigation outlined below as the first to be reported back, and we then discussed its merits as a group.

This investigation used fairly simple statistics, such as correlation, and involved virtually no calculations. The computer did all the work in constructing the scatterplots. The students collected the data from cards and entered it into the database. They were then able to explore what the data said--the reality of the context to them and the relative simplicity of the data meant that they could communicate meaningfully what they found. Even those in the class who had virtually no prior knowledge of basketball (which was the majority, especially among the girls) could understand the conclusions, and actually learned something about the game from hearing the presentation.

The Tabletop software program (Hancock, Kaput, & Goldsmith, 1992) offers an extremely powerful medium in which students can explore multivariable contexts at a fairly young age. The visual representation allows them to deal qualitatively with the interrelationships, and to develop an intuitive understanding of the distinction between correlation and causation. The analysis of a situation that students choose for themselves, and are interested in, provides a framework in which many of the “big ideas” in statistics can begin to develop naturally.

In this very simple investigation, the boys proposed to compare different players in the American NBA, which is now televised in the UK on Channel 4 and on satellite. Their initial idea was that height would be the major factor in the players' performance. They collected raw data on 40 players, choosing 15 guards, 15 forwards, and 10 centers, entered the data into the database, and then investigated the relationship between the various summary statistics that are part of American sport culture. Using an overhead projector with an LCD panel connected to the computer, they presented the conclusions from their investigation to the rest of the class.

They began by explaining the composition of a five-man basketball team (2 guards, 2 forwards, and 1 center) and showed the database they had constructed. In their first slide (see Figure 7), they showed that the correlation between height and average number of points was positive, but weak, and concluded that factors other than height were involved in a player's performance. The grouping of the symbols that show the players' positions was commented on and also in their next slide (not reproduced here), which showed a negative correlation between the height and number of assists.

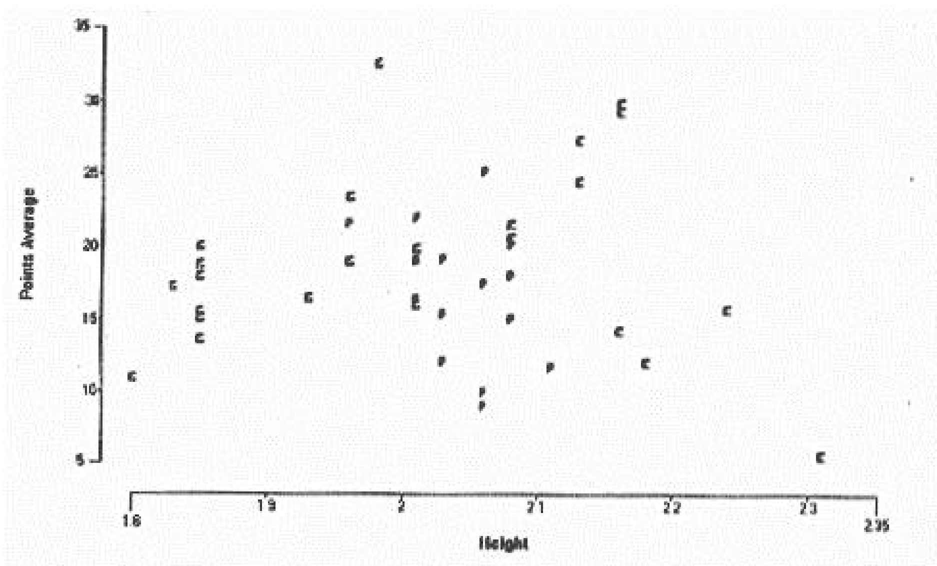

Figure 7: Scatterplot of height and average number of points for each position

Reference to the role of the different positions in a basketball team, and the different height distributions of the different positions, offered a causal explanation of the correlation observed through the relationship of these two quantities to the third quantity (position). Similar treatment of the scatterplots of assists & steals (see Figure 8) and blocks & rebounds developed the understanding of the roles of each of the positions.

I am just starting to explore the possibilities that Tabletop offers for conducting multivariable analyses. Figure 9 shows a scatterplot of lifespan and thorax length from the Fruitfly dataset (Hanley & Shapiro, 1994), which has 5 sets of 25 observations in a designed experiment looking at sexual activity and the lifespan of male fruitflies. Tabletop uses icons, which you can design yourself, to represent individual records. Here, five different icons were used to identify the five sets. The two experimental groups, in which the male fruitflies were supplied with either one or eight virgin females, are labeled with x followed by the number of females. The three control groups, who were supplied with no females or with newly pregnant females, are labeled with a small black square, followed by the number of females. Techniques such as analysis of variance and multiple regression are appropriate for a full investigation of the results of this experiment. However, a good deal of information can be seen informally from the groupings in Figure 9 and in the boxplots in Figure 10, which shows the five groups separately (the visual impact of the scatterplot is considerably greater when viewed in Tabletop, with different colored icons).

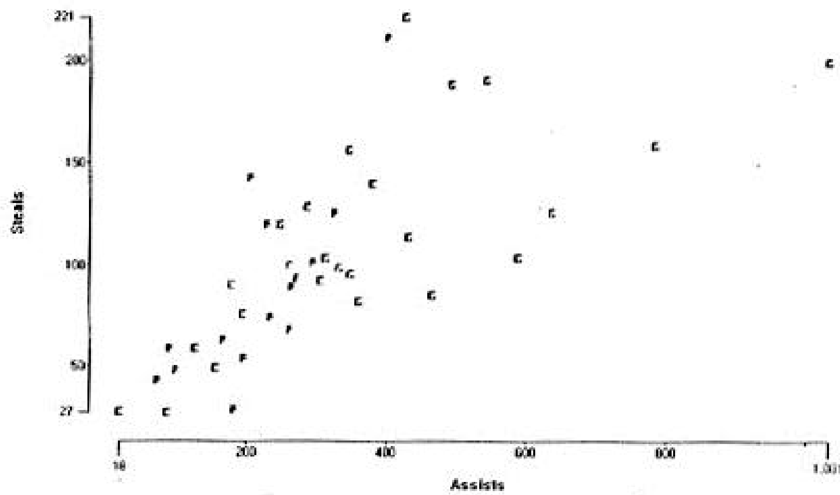

Figure 8: Scatterplot of assists and steals for each position

Discussion of these figures drew out some important ideas. Examining the five groups showed that the two experimental sets had a lower distribution than the corresponding control groups, but there was considerable overlap between all groups; thus, it would not be possible to make an inference about which group a fruitfly came from by examining its lifespan. By examining Figures 9 and 10, the students concluded that the factor being controlled here is not the only influence. This generated debate that centered around the distinction between association, which they could see, and causation, which they had no backgound knowledge to determine the direction of. Interestingly, there was an inherent understanding in this context that the variation between fruitflies within a group was the result of other factors, rather than "errors," which the language of formal regression studies pushes them toward.

UNDERSTANDING VARIABILITY

One of the biggest difficulties I encounter in teaching statistics to students or in discussing the teaching of statistics with mathematics teachers who do not have much statistical background in their training, is the underestimation of the amount of randomness that there is in random events. I see various ways in which technology can help in developing a more accurate intuition concerning randomness, by providing an experience-based frame of reference within which intuition can operate.

We all experience random events--even 11 and 12 year-old students have familiarity with it from their experience in playing games of chance. They do not experience it in any systematic way, however, and misconceptions such as "six is the hardest number to throw" are observed. The process of systematically collecting observations from random events such as throwing dice, or tossing a number of coins, is a fairly tedious and time-consuming one, but I think that the collection of data in a context that the pupils can easily relate to is worthwhile. However, technology can be used in two ways here to take some of the drudgery out of the process.

3. DEVELOPING PROBABILISTIC AND STATISTICAL REASONING

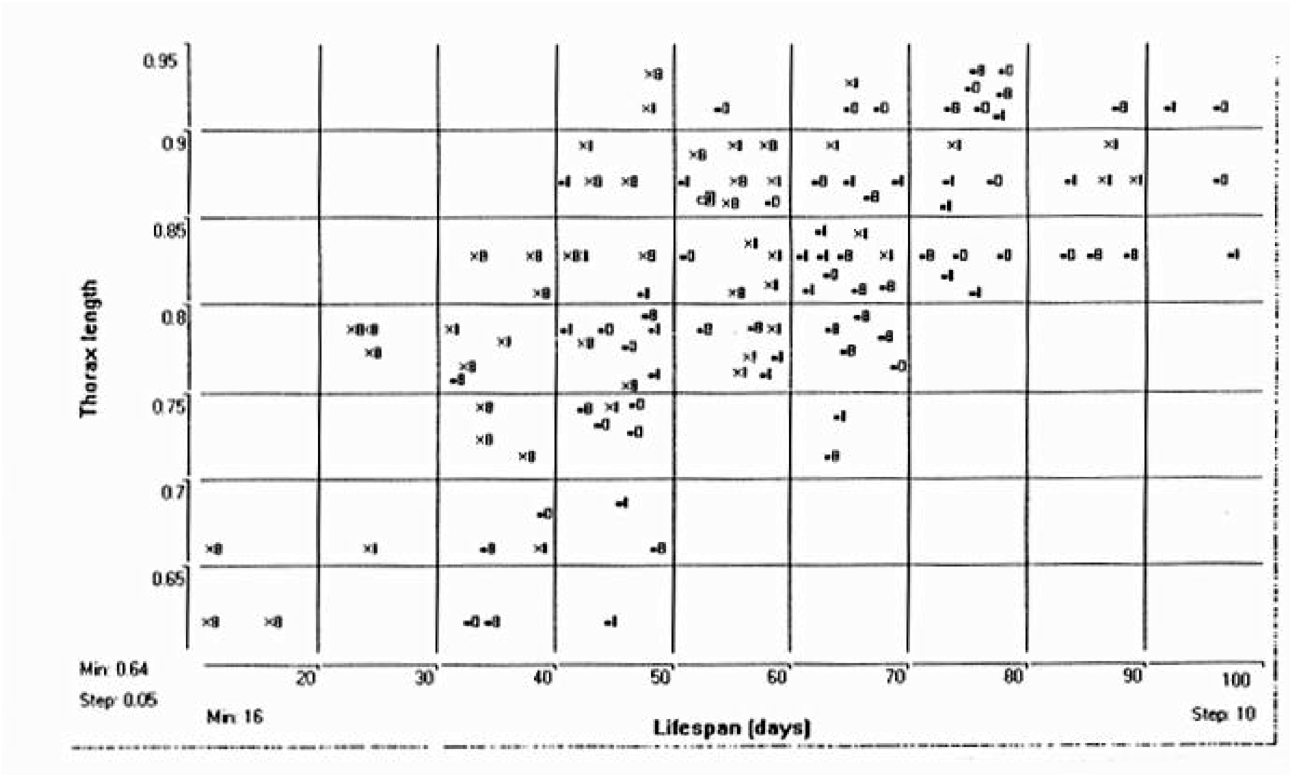

Figure 9: A scatterplot of lifespan and thorax length from the fruitfly dataset

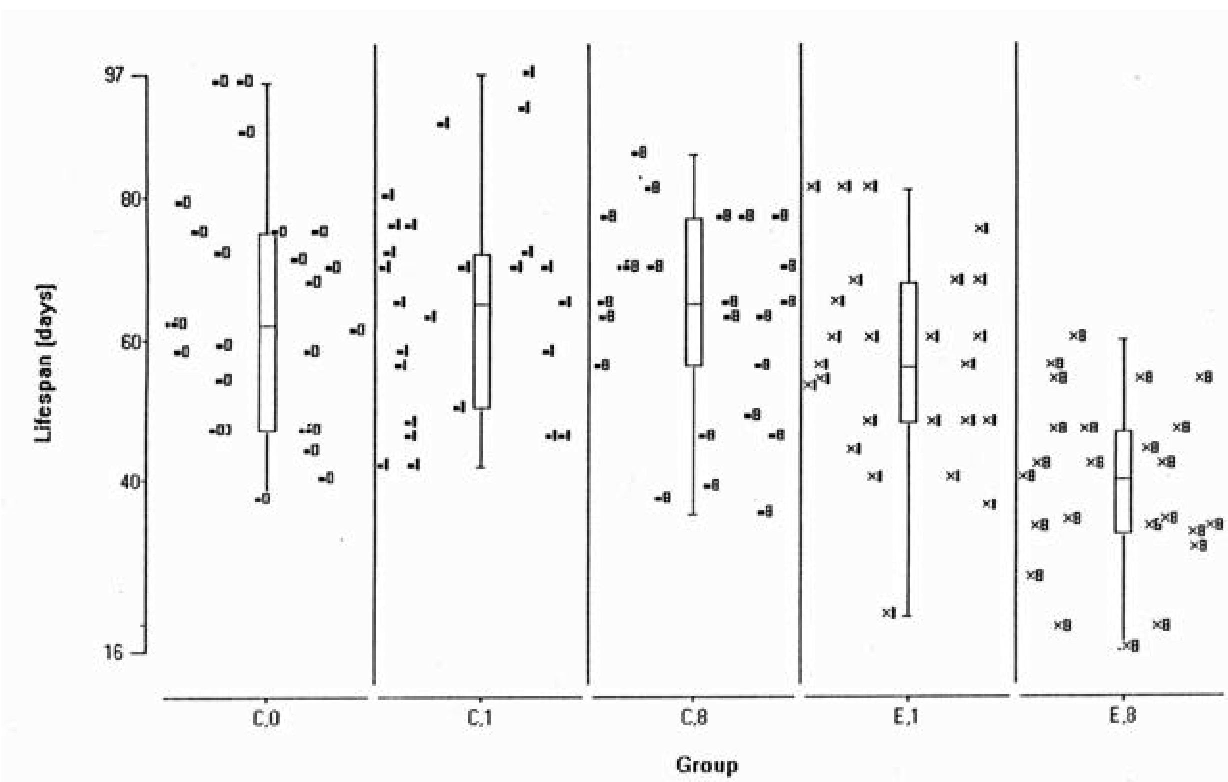

Figure 10: Boxplots of experimental and control groups in the fruitfly dataset

The graphs in Figures 11 and 12 show data that was collected by a Form 1 (Grade 7) class. Each pupil recorded the scores on 100 throws of a die. The throws were actually recorded in order of occurrence so that we could also analyze how often runs of 2, 3, 4, … of the same number happened. The totals for each pupil were entered into a spreadsheet. The students then grouped sets of six pupils together, and then grouped the entire class together. Figure 11 shows the first four pupils' results, and four groups of six results are shown in Figure 12. By being able to see the variation in individual results for a substantial number of cases and also for larger groups, an understanding of the way the proportions observed behave as larger groups are taken can be fostered over a period of time. The various individual student graphs and the graphs of the grouped results are displayed in class and are referred back to at various stages later in the course when we are dealing with variation again.

Consideration of the rate of occurrence of single numbers and runs of 2, 3, 4, … develops an understanding of the problems involved in estimating unknown probabilities--with 100 throws each, the students think they have a lot of data. The theoretical proportions are 83% for singles, 14% for runs of two, 2% for runs of three, 0.4% for runs of four, and in a class of 29 pupils it would be unwise to bet against getting at least one run of five or more. Table 2 shows the distribution of the lengths of the runs obtained by eight students using a fair die; there is considerable variation in the observed distributions collected by the students. The students did not know what the theorectical distribution was for this situation, and the variation in the observed results provided the basis for a lively discussion of what the true proportions actually are.

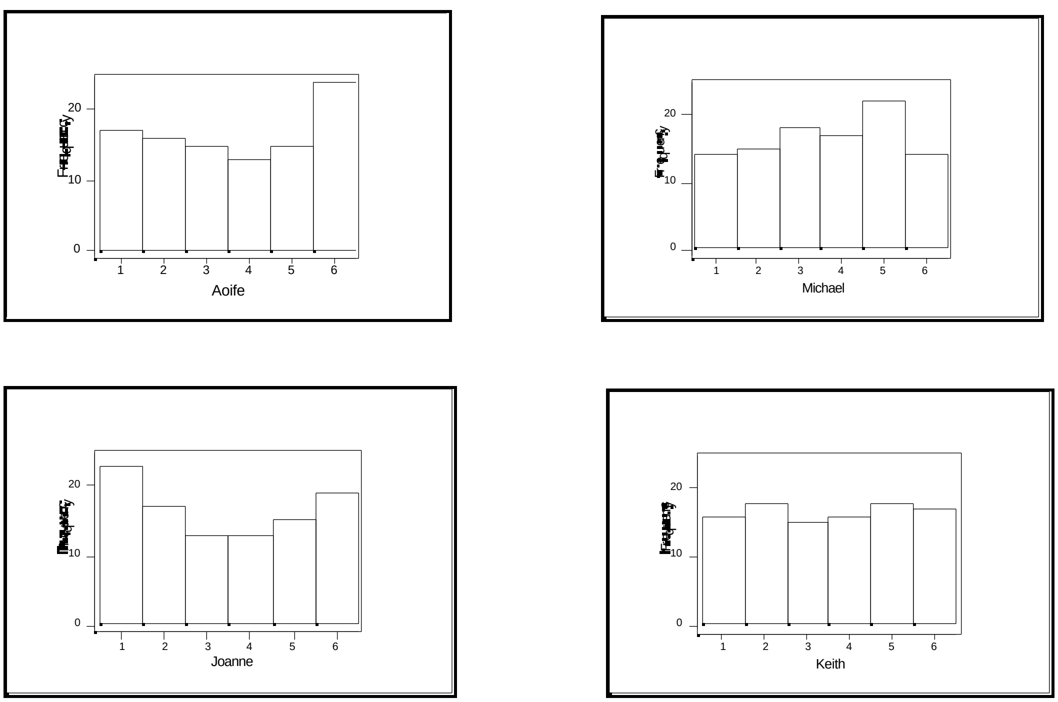

Figure 11: Dice throwing results for four different students

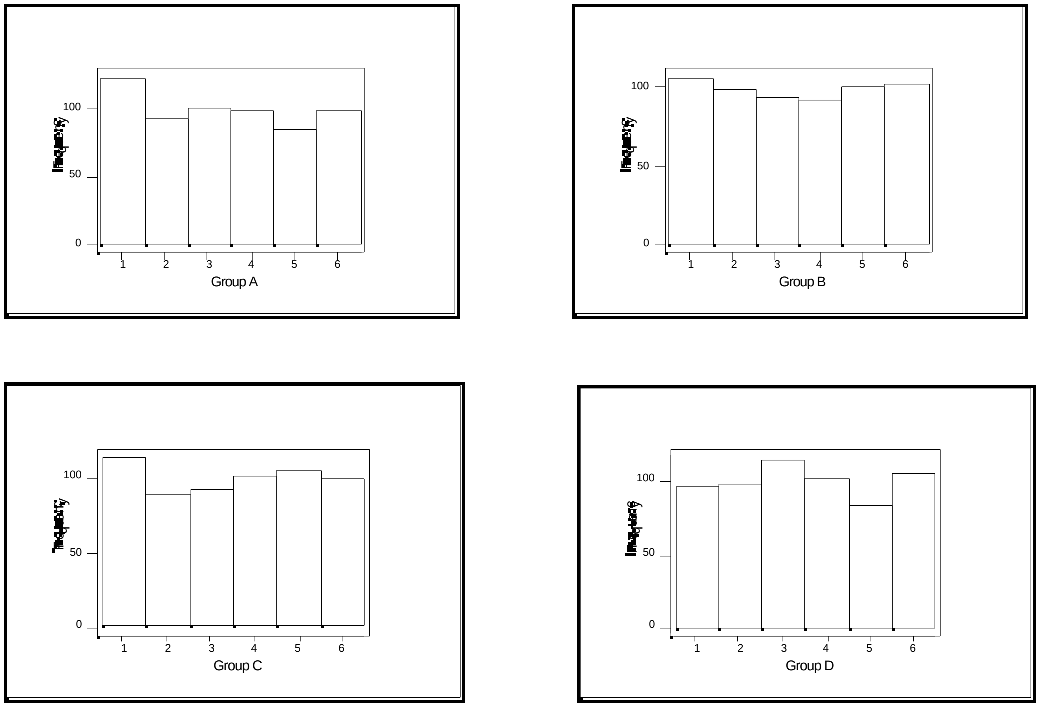

Figure 12: Results of dice throwing for four groups of six students each

Table 2: Lengths of runs observed using a fair die

RUN	1	2	3	4
David	67	15	1	0
Emer	76	9	2	0
Keith	48	18	4	1
Ryan	58	11	2	3
Laura	75	8	3	0
Carole	63	14	3	0
David	79	9	1	0
Jenny	64	12	4	0
TOTAL	**530**	**96**	**20**	**4**

This also works well with a biased die, but I find one advantage of the proportions of lengths of runs is that with guidance the theoretical distribution is easy for the class to derive and check against the distribution of the pooled results. The discussion following this exercise can be lively, as I have already said, and for many pupils the actual results are the main, if not the only, focus of attention. However, some pupils begin to show a qualitative understanding of some of the principles of interval estimation. Certainly, the entire group displays an enhanced grasp of the principle that relative frequencies of observations provide estimates of the probabilities of events occurring.

SIMULATIONS AND OTHER ELECTRONIC AIDS

The process of generating random data is one that can be done very efficiently using an electronic medium such as a graphical calculator or a computer program. However, I find that some students find it hard to accept initially that the medium accurately models the "real thing." They often have poorly formed ideas of how randomness behaves. For example, they might believe that if five heads have appeared in a row, then the law of averages means the next should be a tail, and so forth. When an electronic simulation does not match their intuition, it is more comfortable to think it is doing something different rather than to adjust their intuition. Following up a dice throwing experiment with a dice throwing simulation, and finding similar degrees of variability in the two sets of results, seems to reassure the students. I think that it is time well spent, because the students find it easier to imagine, in the future, how a simulation can be set up if they can think of the concrete cases they have in their own experience. Once they are convinced that the computer's random behavior is similar to the real thing in one context, it ceases to be an issue.

One can use commercially available simulation software; however, it is now relatively easy to generate fairly informative simulations using macros in statistical packages and in spreadsheets without actually having to program. Also, the Discovering Important Statistical Concepts Using Spreadsheets (DISCUS; Hunt & Tyrrell, 1995) materials exist as a series of electronic workbooks in EXCEL, and have some excellent simulations in them that are ready to run! DISCUS also uses EXCEL's data analysis tools and charting facilities to provide very strong visual images of relationships that exist. For instance, students can explore the relationship between the binomial, Poisson, and normal distributions by superimposing distributions in which the student defines the parameters. Tyrrell (1996) provided a fuller description of the material covered and of the style of the materials.

APPLICATIONS OF SAMPLING

Over the past few years, I have been moving away from a fairly theoretical-based delivery of the A-level Statistics course that I teach toward trying to find ways that the students can actually have some experience of what the results mean in practice. As I have argued already in this paper, I believe that such experience informs their intuition; thus, their analysis of a new problem situation is more likely to be accurate.

Initially, this development in style was not based on the use of technology. It was based on the collaborative efforts of a large group in which the group would look at the same dataset in different ways and pool the results, which is similar to the way that I would pool the results of a dice throwing experiment with the younger students.

Sampling, linear regression, and correlation are topics that are encountered early on in most people's statistical experience, whether it be in a mathematical statistics course or in applications of statistics in other subject areas. The least squares line of regression and the product moment correlation coefficient are well defined functionally, and many calculators and computer software packages will generate them after elementary data entry procedures are followed.

I teach an A-level course that has 50% statistics in which these topics appear. I had been concerned that the students' perceptions of the regression line seemed almost to extend to a belief that this was **the** underlying relationship, and that the line and predicted values generated by the line of regression would be given to very high degrees of accuracy. These would be unwarranted even due to the accuracy to which the data had been recorded, before any consideration of the variability of the line due to sampling. I had also been concerned about the students' grasp of confidence intervals and particularly how dependent they are on the data used, and we were going to be looking at different sampling methods. We had already spent some time early in the course working with boxplots and had used them to make comparisons between datasets with different centers and spreads.

Burghes (1994) contained a dataset that contains a variety of information on 200 trees on a piece of land that the owner wishes to sell. It is in a chapter dealing with data collection, and the activity suggested is for students to choose one of a number of sampling methods and construct an estimate of the average value of the various quantities, such as value, age, girth, and the proportions of different types of trees. The class undertook this activity, and we then pooled all the results and constructed boxplots of the data resulting from the various (point) estimates of these quantities using random, systematic, and stratified samples. From the class 20 students, six or seven estimates were obtained for each sampling strategy.

These boxplots provided considerable insights into a number of different and difficult concepts, which I will expand on below.

SAMPLING TYPES

The nature of stratified sampling is illustrated by Figure 13, in which the proportion of oaks in each sample remained the same, because that was how the stratification had been constructed. The perfectly consistent prediction (of the true proportion) provided a focus for what makes a good estimator. In the same diagram, the contrast between the profiles of the random and systematic sampling estimates opened up some worthwhile discussion as to why the systematic samples provided much more consistent estimates than the random samples in this case and whether we could expect that in all cases.

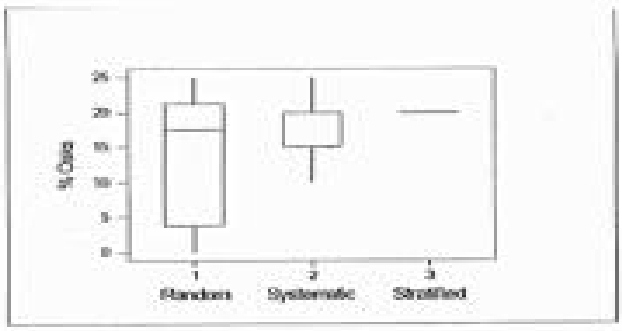

Figure 13: Example of stratified sampling

Figure 14 shows the estimates of the monetary value of the trees from the different sample types, which led to a discussion questioning whether the size of our samples was sufficient to allow us to make strong statements concerning the relative merits and demerits of different sampling methods. Because the students had to do the work in producing the data from samples, they showed a greater understanding than previous groups of the “costs” involved in improving the quality of your conclusions by considering more data.

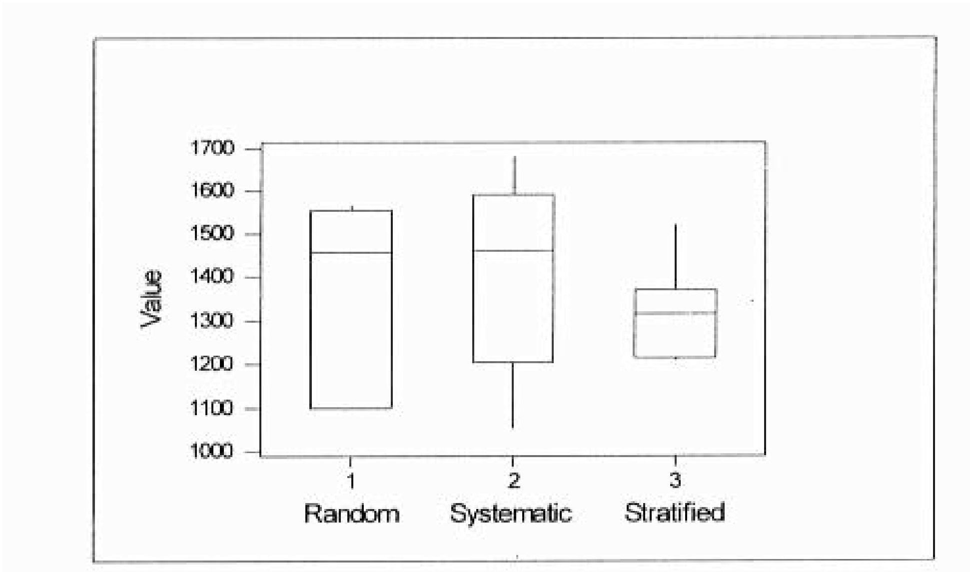

Figure 14: Boxplots of estimates of the monetary value of the trees

INTERVAL ESTIMATION

The ideas of interval estimation were able to be developed quite naturally from the experience of generating a number of different point estimates of the same quantity and finding that they did not always produce the same value for the estimate and that the consistency of the estimate was affected by a wide range of factors, such as the size of sample used, the method of constructing the sample, the underlying variability of the quantity, and so forth. This led to interesting investigations of other datasets in an attempt to quantify some of these effects.

The process of repeating a number of samples using the same procedure to generate them, but obtaining different datasets each time and sometimes quite different estimated values, meant that the students had an experience to draw on, which informed their intuition when dealing with subsequent situations involving sampling. Instead of a set of abstract rules (e.g., that the variance of estimated values varies inversely with the sample size) that previous classes could learn and apply, often very successfully, this group began to appreciate intuitively, based on experience, how the consistency of the estimates would vary.

REGRESSION AND CORRELATION

This led to examining the dependence on the data used in other circumstances, in particular in linear regression and correlation. Table 3 shows the body and heart masses of 14, 10-month old, male mice. We examined the regression lines and the predicted heart masses for certain body masses that would be generated by samples of the dataset.

Each of the data points was discarded in turn. The values of the correlation coefficient, the coefficients of the equation of the line of regression of heart mass on body mass, and the predicted values of heart mass for body masses of 20, 35 and 55 grams were computed. Figure 15 shows the predicted values obtained in boxplots. Much greater variation arose when smaller samples (e.g., using 10 out of 14 mice) were chosen and when other datasets were used in which the correlation was not as strong.

Table 3: Dataset of body and heart mass of 14 mice

Mouse	Body mass (g)	Heart mass (mg)
A	27	118
B	30	136
C	37	156
D	38	150
E	32	140
F	36	155
G	32	157
H	32	114
I	38	144
J	42	159
K	36	149
L	44	170
M	33	131
N	38	160

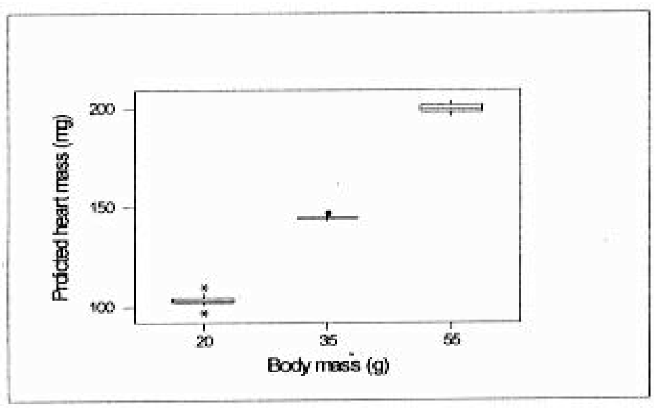

Figure 15: Predicted values shown in boxplots

Again, a number of important issues arose from this exercise:

- The regression line was appreciated more as being an estimate of the underlying trend.
- The regression lines based on various samples were seen to be diverging as you move from the center of the x-values, and the boxplot above shows clearly the greater consistency of predicted values close to the middle of the range.
- The problems associated with extrapolation take on a new dimension. Not only may the existing fairly strong linear relationship not continue, but even if it does the predicted values are increasingly unreliable.

Upon reflection, I was pleased at how coherently the different aspects of this work fit together and indeed even provided reinforcement for the other concepts. I was also particularly pleased at how the students' intuitive understanding of some quite difficult and subtle concepts seemed to be more secure because it was experientially-based to some degree rather than purely learned by formal mathematical principles (theorem and proof), even though these are still an essential part of statistics.

I then started to examine whether technology could help in the production of the data as well as with the analysis and visual representation. The diagrams below were produced using the student version of Minitab for Windows (Version 9). The small routines were written for Minitab (called "executable macros"), although I believe the principles can be adapted quite easily for other software and hardware.

I will deal explicitly with three cases, but the principles and methods involved are applicable in a variety of other situations and also reinforce a number of other ideas in passing, which I will try to touch on briefly. I have based them on data that was randomly generated within the macros used, but I would recommend working with real, large datasets when possible.

DISTRIBUTION OF SAMPLE MEANS

The variance of the mean of a sample of size *n* from a population of variance σ^2 is σ^2/n. Many students have an intuitive feeling that a larger set of observations should provide a better estimate, without any firm grasp of the criteria on which this could be judged. Indeed, it is very difficult for them to resolve the interrelating nature of the distribution of the population, the distribution of the sample, and the sampling distribution of a statistic such as $\bar{x}$. The macro developed here uses the following procedure: (1) it asks the user to provide the mean and standard deviation of the population and the sample size to be used; (2) it generates 40 columns of data each with the requested sample size (limited to 80 when 40 sets are used, because Minitab student worksheets are limited to 3,500 cells); and (3) it then computes the mean of each of the 40 columns, before drawing the graph of the sample means as a boxplot. Using 40 pieces of data to draw the boxplot allows one to get a good idea of the variability of the distribution of the sample means in each case. By systematically varying the standard deviation and the sample size, the student can build up an experience-based intuition of the behavior of the variability of the sample mean. Figure 16 shows the cases in which the standard deviation and sample size were 5 and 80, 5 and 10, and 15 and 10, respectively. The population mean in all cases was 34. Apart from the changing behavior of the variability of the sample mean, there are a number of important other points that can be made: there are 40 point estimates of the population mean used in each boxplot, and the basis of confidence intervals as estimators, which are more informative than simple point estimates, can be seen comparatively easily. In all cases, the point estimates are centered around the true value of the parameter. With a population standard deviation of 5 and a large sample of 80 observations, the sample means gave estimates of the population mean that are consistently very close to the true value, whereas with a larger population standard deviation and smaller sample this is not so.

For this example, all three values for mean, standard deviation, and sample size were entered using the keyboard while the macro was running. It can usefully be altered for teaching purposes so that the teacher has already specified the mean and the standard deviation in advance, and the students are then looking at a real estimation problem.

The data used here were simulated observations taken from a normal distribution, but a simple alteration to the macros would allow samples to be drawn from any population listed as a column in a worksheet. The effects of sampling with and without replacement can be investigated. Note that the variance of the sample mean is reduced when sampling is taken without replacement, being multiplied by a factor $(N-n)/(N-1)$ for a sample size *n* from a population of size *N*.

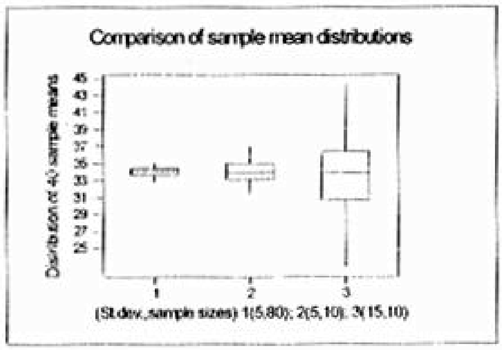

Figure 16: Comparing sample mean distributions

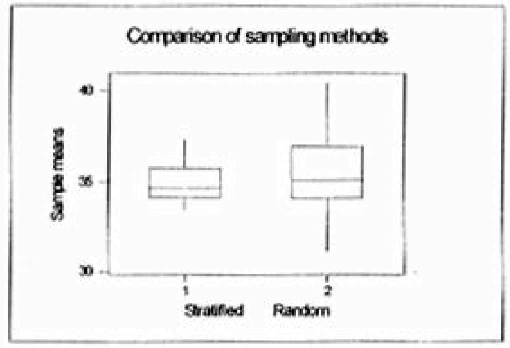

Figure 17: Comparing sampling methods

Stratified versus random samples

The effect of stratified sampling where a population consists of a number of identifiable groups (strata) rather than one homogeneous group is to greatly increase the consistency of the estimator, by removing one of the sources of variance (i.e., the differing number of observations from each strata). The macro used here performs the following procedures: it generates sets of 40 observations from normal distributions with means of 25, 30, 35, 40, and 45 each with a standard deviation of 5; it then repeatedly samples 4 observations from each set and also 20 observations from the full set of 200, in both cases without replacement; as before, the sample means of each of these sets are computed and boxplots are drawn to compare the consistency of the estimators. Figure 17 shows one outcome of this process.

The macro can be altered easily to work with real stratified populations, to accept parameter values for each strata as inputs using the keyboard, or for the teacher to assign values for the strata parameters in advance, placing greater emphasis again on the realities of interval estimation rather than point estimation.

These two investigations also give students some extra experience of using boxplots to make comparisons between distributions.

CORRELATION COEFFICIENTS FOR SMALL SAMPLE SIZES

I find it fairly easy to justify to a class that the correlation coefficients for small datasets need to be high before you can be reasonably confident that there is any underlying relationship, but much harder to justify the distribution results listed in their sets of tables. Running a simulation generating sets of points, under the bivariate normal hypothesis, and calculating the correlation coefficients, allowed us to build up a sampling distribution, from which the origin of the critical values listed could be seen. Two examples of the sampling distributions obtained are shown in Figures 18 and 19.

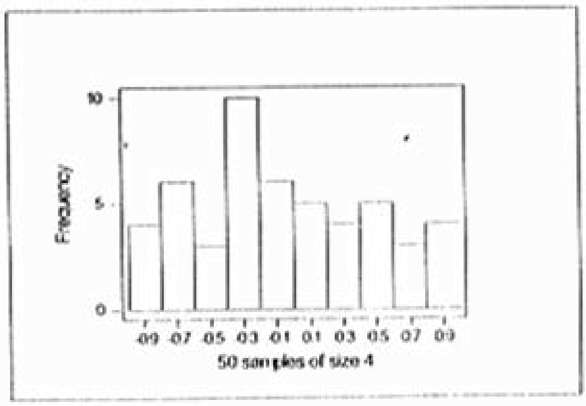

Figure 18: Sampling distributions (n = 4)

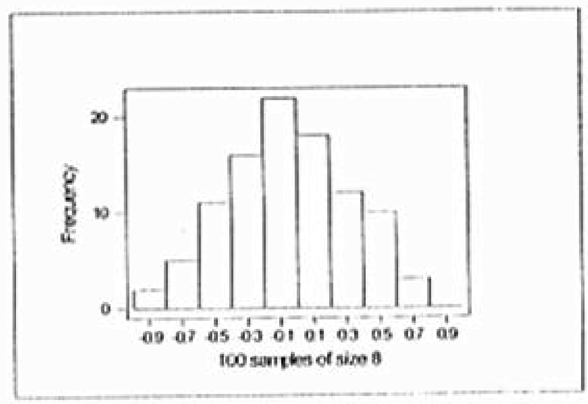

Figure 19: Sampling distributions (n = 8)

CONCLUSION

Much of probability and statistics requires a different type of thinking than mathematics, yet most of the teaching at the secondary level is in mathematics. Also, these teacher often have no significant training in the areas of probability and statistics. The earlier that students are shown the ways that data behave, rather than learning the “mathematical rules” that govern certain aspects of that behavior, the more likely we are to produce students who are genuinely at ease in dealing with uncertainty. I believe that the use of technology and real datasets offer us greater possibilities of doing this well, but further research is needed to determine the extent of the effect. Other papers in these proceedings (e.g., Lajoie, 1997) also argue that datasets should be interesting and relevant to the students’ lives, which helps to motivate them and to show how statistics is used to make everyday decisions on both major and minor issues.

Some specific questions for future research raised here include:

- Does the study of real datasets in the early formative years improve students’ ability to interpret formal statistical results later on?
- Does the cooperative study and reporting of statistical investigations make students better suited to the employment market than an undiluted diet of competitive, individual assessment?
- Does the use of simulations and the sort of sampling investigations described in this paper help students to have a more accurate intuition of what is happening in other stochastical situations?

REFERENCES

Burghes, D. (1994). *Statistics: AEB mathematics for AS and A-level*. Exeter, England: Heinemann.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337-364.

Hanley, J. A., & Shapiro, S. H. (1994). Sexual activity and the lifespan of male fruitflies: A dataset that gets attention. *Journal of Statistics Education, 2*(1). http://www.stat.ncsu.edu/info/jse/homepage.html.

Hunt, D. N., & Tyrrell, S. (1995). *DISCUS*. Coventry, England: Coventry University Enterprise Ltd.

Tyrrell, S. (1996). Statistics using spreadsheets. *MicroMath, 12*(2), 32-36.

3. DEVELOPING PROBABILISTIC AND STATISTICAL REASONING