22. COMPUTER PACKAGES AS A SUBSTITUTE FOR STATISTICAL TRAINING?

Michael Wood
University of Portsmouth Business School

INTRODUCTION

Proposals are made for coping with the problems of teaching statistics to managers, and to students of management in higher education. The problems in question concern the fact that teaching statistics in these contexts is difficult and often ineffective: Numerous anecdotes indicate that these clients often neither like nor understand the statistics they are taught (see Wood & Preece, 1992, for a discussion of some of the problems this causes).

For these clients, the discipline of statistics is a means to an end of being a better manager. Understanding statistics is not an end in its own right, as it might be for students on a degree course in statistics. It is for this reason that I have used the term "training" instead of "education" in the title. This is not to say that the purpose of learning statistics is simply to solve a particular problem or set of problems: Long-term utilitarians will learn ideas now because they think they may be useful in the years to come. However, if the subject is misunderstood, or avoided because it is disliked, it is of little practical use.

This paper focuses on statistics and management, but I believe that very similar comments apply, for example, to statistics and medicine (Altman & Bland, 1991) and to mathematical modelling and management. In short, this paper addresses the problem of helping any novices use mathematical methods of any kind.

The usual approaches to the difficulties of teaching statistics are based on advice such as practice more, think harder, get a better teacher, or even worse, become more intelligent. I think the problem is too difficult to be solved simply by these methods. Accordingly, this paper concentrates on a potentially more powerful strategy: that is, changing the methods and the technology used to make statistical analysis easier. The principles proposed here are not fully tested; however, I will illustrate them using specific examples, most of which I have used on an informal basis.

THREE HYPOTHETICAL SCENARIOS: A THOUGHT EXPERIMENT

I will start with a thought experiment to compare three (similar) groups of students subjected to very different "treatments." Let us say that the students are groups of managers taking a course on statistical process control (SPC).

Treatment 1. Training plus (statistics) package

This is the conventional treatment: a standard course, plus another course on the use of the statistics package to implement the techniques taught once the students have "understood" them.

Treatment 2. Training by (training) package

This group is taught exactly the same as the first group (i.e., they are exposed to the theory and to the use of the package), but the teaching is done by an intelligent, multimedia computer system. If such a system automates the best teaching practice and includes some features that the best of human teachers could not incorporate (e.g., providing help 24 hours a day), it must be better than a conventional training course.

Treatment 3. (Statistics) package as a substitute for training

The third group does not attend a course at all. Instead, they are provided with a package that assists them with statistical analysis as and when they need it. The package would be largely computer-based, but might also incorporate paper-based elements such as instructions for drawing simple diagrams and statistical tables. The essential feature of the package is that it is designed to be used on the job when required to solve problems. It may incorporate appropriate "intelligent" front and back ends to interface with a novice user of statistics; that is, it would provide guidance on the correct statistical approach to use (the front end) and on the correct interpretation of the answers (the back end).

How do these three groups compare? If the training package used in Treatment 2 is as good as, or better than, the best human teacher, then by definition Treatment 2 is preferable to Treatment 1. A similar logic indicates that Treatment 3 is preferable to Treatment 2 because both satisfy the goal of enabling "students" to perform and understand statistical analyses, but Treatment 3 cuts out the necessity for training and is thus more efficient in terms of students' time. It then follows that Treatment 3 is the best one.

In practice, limitations in the effectiveness of the software mean that these conclusions are not fully justified. Training software is not a complete and adequate replacement for a human teacher; thus, a combination of Treatments 1 and 2 (human teacher backed by training packages in areas where these are effective) is likely to be better than either in isolation.

In a similar vein, Treatment 3 is impossible because no packages are sufficiently "intelligent" to enable a novice to do statistics properly. There are two main reasons for this: (1) the lack of a suitable conceptual framework for communicating with the user, and (2) the fact that users are likely to lack a suitable image of what the package will do. These are discussed below.

Difficulties with Treatment 3

In an earlier paper (Wood, 1989), I illustrated these difficulties by means of a largely unsuccessful attempt to build and use an expert system to enable novices to use standard statistical distributions (see Appendix 1). The expert system had no difficulty with the arithmetic nor with implementing "general rules," such as the fact that the Poisson distribution is a reasonable approximation to the binomial when the "sample size" is large and the "probability of success" is small. However, potential users did not understand what "sample size" and the "probability of success" meant. Obviously, these concepts could be explained using a "help" key, but this can only be done to a limited extent because the package cannot explain all the ways in which the notion of a "sample" can be applied to a *new* situation. Users simply have to *understand* the concept in a fairly deep and flexible sense. Furthermore, users had no image of what a statistical distribution was nor of the types of situations that the standard distributions will model adequately; thus, they did not know when the package was likely to be useful.

A further difficulty is the issue of trust in answers provided and the chance of an unrecognized error occurring. The package may give the "answer" and imply that the answer is correct, but in many cases answers are not correct and sensible users will want a justification of the answer so that it makes sense to them. This is clearly connected with the interpretation of the output and with the user's concept of what the package can be used for--users who have reason to believe the answers produced by a package are much more likely to interpret the output meaningfully and to understand when the package is likely to be useful.

In principle, there are three types of justification. First, the status and consequent authority of the statistical package might satisfy some users [e.g., SPSS (Norusis, 1993) gives this answer so it must be right]. The second alternative is for the user to understand the algorithms used by a package; however, in practice this is unrealistic even for quite "simple" algorithms, such as those for computing normal probabilities. The third alternative is for the user to check the output against different criteria. For example, a linear regression equation can be checked by plotting it with the data, a significance test routine can be evaluated by examining whether it behaves sensibly with large and small samples and with data showing obvious patterns and data showing no such patterns.

How does this affect the comparison between Treatments 2 and 3? Building an intelligent training system presupposes the existence of a computer model of the content of the training. It also presupposes that a model of the learning process and a model of an individual learners exist. This means that the statistical content of the packages used in Treatment 2 can never be more--and will almost certainly be less--than that of the intelligent packages used for Treatment 3. This merely reinforces the superiority of Treatment 3, and suggests that Treatment 2, in its pure form, is never a sensible option.

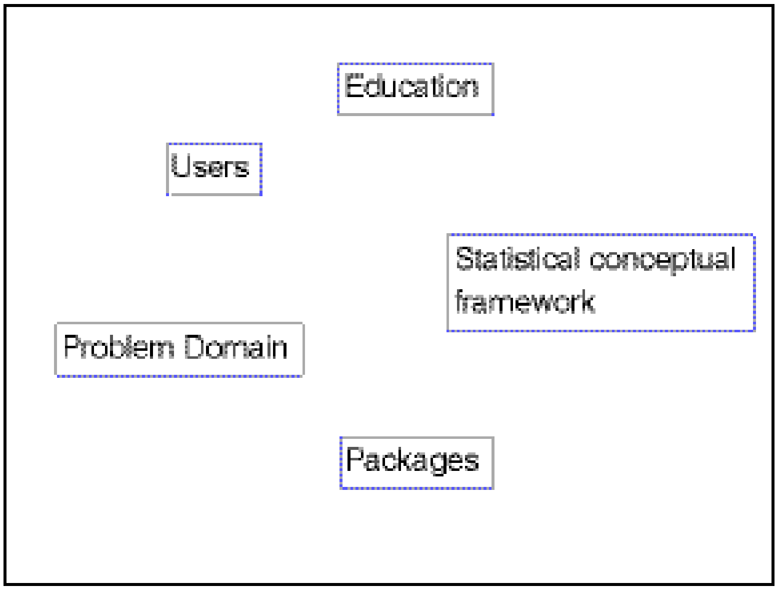
However, Treatment 3 is not possible in its pure form either, because users of intelligent statistical packages need some education in underlying concepts. In particular, users need the conceptual background to be able to understand the input requirements and to interpret the output, to understand what the package and the statistics it implements "does,” and to be able to check if they are using the package correctly so that they can use it with confidence.

This means that the best treatment is a combination of Treatments 1 and 3. The intelligent training systems proposed in Treatment 2 are ruled out because whenever they are viable the expertise must have been modeled; thus, the more efficient Treatment 3 is viable. Treatment 3 is the most efficient and can be considered the baseline, but the education of the users in order to use the statistics package appropriately (Treatment 1) must be acknowledged. Accordingly, this will be called Treatment 3-1.

As technology progresses, the likelihood is that the power of intelligent statistical packages will increase, and the consequent need for education will decline. We will need to teach fewer or less difficult concepts to achieve the same results. However, I cannot envisage a state in which there will be no necessity for the education of users.

Treatment 3-1: A paradigm shift

Treatment 3-1 represents a paradigm shift with corollaries that are more far reaching than may be apparent at first sight. The beneficiaries of Treatment 3 are *users*, not students. This may indicate a shift in more than terminology--academics using packages such as SPSS to produce *p* values to legitimize their research are users in the sense of Treatment 3, but are not students because they do not take courses. Similarly, the treatment uses a *package* instead of a course. Treatment 3-1 incorporates the necessity for user education. But the Treatment 1 question "How should we teach it?" has changed to "How can we make it easier?" or "How can we avoid it?" for Treatment 3-1. The elements of the system are represented in Figure 1.



1

Figure 1: Elements of the system

Under Treatment 1, any problems or mismatches are dealt with by manipulating just one element of this system--education. Package design is seen as a separate issue, because this depends just on the statistical framework. The change of emphasis in Treatment 3-1 suggests that the packages can also be changed if this helps the system. Treatment 3-1 also raises the possibility of treating the statistical conceptual framework as a variable. The traditional educational paradigm (Treatment 1) implicitly treats the statistical content as a given: The job of education is to introduce students to as much of the established knowledge as possible. With the more utilitarian perspective of Treatment 3-1, it is natural to consider the possibility of changing the statistical framework itself. If the customers do not understand a concept or technique, try something more user-friendly. With three elements of the system considered as variable, instead of just one, the optimum configuration should be better.

For designers of commercial software packages this attitude would be accepted without thought. Academics, however, tend to have an inertia because they are seen, in some sense, as "absolute." There is a strong argument that this attitude is unreasonable. The development of science and other branches of human knowledge are very much influenced by fashions and social and technological pressures of various kinds (Kuhn, 1970). From this perspective, Treatment 3-1 is simply a reflection of changing social and technological circumstances. These conclusions suggest two principles with specific practical corollaries--the simplicity principle and the black box principle.

The simplicity principle

It was argued above that, regardless of the degree of intelligence incorporated into a statistical package, it was essential that users understand the conceptual descriptions of the input and output. In addition, users should have an understanding of the procedures implemented by the package so that they can avoid errors and have confidence in their answers. It follows that the easier the statistical concepts and procedures are, the better--the simplicity in question here is conceptual simplicity from the perspective of users. It is important to ensure that users can understand what the package is doing and what the answers mean.

It is often surprisingly easy to devise simpler methods and concepts. Wood (1995) suggested that part of the SPC (statistical process control) syllabus for the subjects of this thought experiment could be simplified using a resampling procedure instead of conventional formulae based on probability distributions (see Appendix 2). The advantages are that the method is more transparent; that is, users can see what is happening and how it works *without using the usual mathematical models of probability distributions*. It is a more general method because one procedure replaces a range of different models, which makes it conceptually easier for users. It is also, perhaps surprisingly, more rigorous than the conventional approach in many situations because the conventional formulae depend on crude assumptions (e.g., of normal distributions) and make rough approximations (e.g., the normal approximation to the binomial when p is small). There are also other possibilities (e.g., an alternative framework for control charts is suggested in Wood, 1995).

In a more general context, resampling methods can be used for statistical inference (Jones, Lipson, & Phillips, 1994; Kennedy & Schumacher, 1993; Noreen, 1989; Ricketts & Berry, 1994), and non-parametric methods in general provide a simpler but effective approach to many problems.

The black box principle

Sometimes simplification is not possible, but the underlying ideas may be too complex for the intended users. In these circumstances, the computer package has to be treated as a black box; that is, users do not look inside and do not try to follow the algorithms used because even if they did it would not help them. Academics using statistical packages for calculating p values are often in this situation.

There is a potential problem here if the black box is used incorrectly, if the output is misinterpreted, or if key assumptions are not recognized. Two approaches to reducing the severity of this problem can be identifed.

1. User-friendly input and output. Process capability (to return to the SPC course) is usually measured by capability indices based on a normal model that gives a value of approximately 1 for a process producing approximately .2% defectives (the conventional "three sigma" percentiles of the normal distribution), and a value of about .7 for a process producing approximately 5% defectives. This statistic, which would be the output of a package for calculating capability indices, is difficult, if not impossible, for users who do not understand the underlying normal model to interpret accurately. (It even creates difficulties for expert statisticians if the normal model is not realistic.)

The answer to this problem is obvious: redefine the index to make it mesh with the user's frame of reference. For example, the package could produce an estimated "parts per million" defective which is far easier to interpret.

For the input, the normal distribution can easily be rewritten so that it does not depend on the standard deviation. The standard deviation is the point at which many novices say they start to lose contact with statistics courses. The difficulty is that the formula for calculating it is relatively complex, and it is difficult to interpret the value of the standard deviation in a straightforward manner. Rewriting the normal tables (or the equivalent computer package) so that the input for the spread of the distribution is based on percentile statistics such as the quartiles or semi-interquartile range is relatively trivial (see Appendix 3). The advantage is that users need to understand percentiles (implicitly at least) to make sense of the output, so phrasing the input in similar terms avoids an unnecessary technicality and thus simplifies the education process. For user groups like the SPC group, "sigma-free statistics" would be an excellent idea.

An alternative to this would be to use a Help screen to explain to users about the standard deviation or capability indices. However, in terms of saving the user time, reducing the likelihood of errors, and increasing the user's confidence, there is clearly no merit in this if the framework embodying the new concepts is of equal or greater power (in the epistemological, not the statistical, sense).

2. Experimentation to compare results with intuitions and to develop intuitions about how the black box works. "Visual interactive modelling" is an approach to modeling in operational research that involves "meaningful pictures and easy interactions to stimulate creativity and insight..." (Elder, 1991, p. 9). It is claimed that these models lead to a number of advantages over nonvisual and noninteractive models. The basic principle is that users can *see* numerical results and graphs recalculated *immediately* after a change is made. Many of the advantages claimed are relevant to novices using statistics. For example, visual interactive modeling is said to help clients *trust* a solution because they can *see* it working on the screen and can check that the solution changes when the input data is changed; for similar reasons, it is likely to help clients spot obvious *mistakes* in models.

A spreadsheet set up to enable users to experiment with the normal distribution provides a good example (Appendix 4; Wood, 1992). For example, the user can change the standard deviation--or another measure of the spread--and watch the shape of the curve change. This should either confirm the user's intuition of the standard deviation in terms of the "width" of the graph, or alert him/her to a misconception. If the mean or the measure of spread was entered incorrectly, this would result in an obviously incorrect frequency diagram, and (with luck) any such mistake would be noticed and corrected. Similarly, a regression routine that allows the user to change the data and see the effect this has on the graph has obvious advantages in terms of enabling users to acquire an intuitive appreciation of the technique.

Visual interactive modeling is, in effect, what the designers of any interactive, graphics-based computer systems strive for. And it is a more complex goal than might be apparent on the surface. Obviously, a visual interactive system should be easy to use and the output should be easy to interpret; it should respond quickly to changes in the data; and it should also respond in such a way that the user can see how *changes* in the input *change* the output.

There is a problem in that the novice may not appreciate how to use a visual interactive model for maximum effect. Some novices will simply key in the data, look at the result, and leave it at that. *The process of experimenting to see how the model works is likely to require encouragement or education, which is an important aspect of Treatment 3-1.*

CONCLUSIONS

This paper focused on students, or other users, whose interest in statistics is purely utilitarian; that is, they need it as a tool for engineering, medicine, business, education, or some other field. Five suggestions were made:

1. A change of emphasis when designing courses and software from teaching the standard concepts and techniques of statistics to enabling students to reason statistically with confidence and without error. The emphasis shifts from methods of teaching "students" to the design of suitable packages for "users," who will still need education in the background concepts.

2. This principle means that the framework of statistical concepts and techniques behind what students learn and the software they use should be treated as a variable and designed to make the cognitive system as user-friendly and powerful as possible.

3. This statistical framework should be as simple as possible. If a simpler framework is (almost) as powerful as a more complex one from the perspective of its likely users, then the simpler framework should be used instead of the more complex one (and not simply as an introduction to it). This means, for example, that methods based on simulation or resampling, and nonparametric methods, are likely to be preferred.

4. Where the technical level is likely to be daunting to students (i.e., when the previous suggestion is an insufficiently powerful principle), the software should be treated as a black box: no attempt should be made to get students to understand the algorithms but instead the students should be taught to develop an understanding of the role of the black box by experimenting with different inputs and outputs. This is likely to be a generalizable and learnable skill. It is also one that software packages can encourage or enable.

5. The concepts in terms of which the inputs and outputs of a black box are phrased should be designed to be as simple, user-friendly, and as few in number as possible. For example, a black box for process capability assessment that takes the raw data as input and gives an estimated parts per million defective as output, is likely to be far more useful than one that takes the mean and standard deviation of the data as input and gives as output a capability index whose interpretation requires considerable expertise. These concepts must be understood thoroughly by users.

REFERENCES

Altman, D. G., & Bland, M. J. (1991). Improving doctors' understanding of statistics. *Journal of the Royal Statistical Society, 154*, (Series A), 223-267.

Elder, M. D. (1991). Visual interactive modelling. In A. G. Munford & T. C. Bailey (Eds.), *Operational research tutorial papers* (pp. 1-16). Birmingham: Operational Research Society.

Gunter, B. (1991, December). Bootstrapping: How to make something from almost nothing and get statistically valid answers--Part 1: Brave New World. *Quality Progress*, 97-103.

Jones, P., Lipson, K., & Phillips, B. (1994). A role for computer intensive methods in introducing statistical inference. In L. Brunelli & G. Cicchitelli (Eds.), *Proceedings of the first scientific meeting of the IASE* (pp. 1255-1263). University of Perugia: IASE.

Kennedy, K., & Schumacher, P. (1993). The introduction of bootstrapping techniques for finding confidence intervals using simulation. *Journal of Computers in Mathematics and Science Teaching, 12*(1), 83-92.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Noreen, E. W. (1989). *Computer intensive methods for testing hypotheses*. Chichester: Wiley.

Norusis, M. J. (1993). *SPSS for Windows base system user's guide release 6.0.* Chicago: SPSS.

Ricketts, C., & Berry, J. (1994). Teaching statistics through resampling. *Teaching Statistics, 16*(2), 41-44.

Wood, M. (1989). Expert systems, expert tutors and training in elementary statistics. In N. Shadbolt (Ed.), *Research and development in expert systems, VI* (pp. 195-206). Cambridge: Cambridge University Press.

Wood, M. (1992). Using spreadsheets to make statistics easier for novices. *Computers Education*, *19*, 229-235.

Wood, M. (1995). Three suggestions for improving control charting procedures. *International Journal of Quality of Reliability Management*, *12*(5), 61-74.

Wood, M., & Preece, D. (1992). Using quality measures: Practice, problems and possibilities. *International Journal of Quality and Reliability Management, 9*(7), 42-53.

APPENDIX 1: AN "EXPERT" SYSTEM FOR STATISTICAL DISTRIBUTIONS

The expert system in question, PROP, is a simple knowledge base built with the shell CRYSTAL. The students were enrolled in a statistics class for business studies in a college. It was obviously not feasible to build a system that would cover the entire area of applying statistics to real business situations, but a system that would cover the following topics seemed viable:

- Choosing appropriate measures of location and spread (e.g., mean, median, standard deviation).
- Calculating measures of location and spread.
- Choosing a suitable distribution to model a situation (e.g., normal, binomial).
- Calculating proportions of a population in a specified range (using the appropriate distribution).

Whether such a system deserves the title "expert" is debatable. However, CRYSTAL is an excellent tool for developing such a system.

The main problems with the prototype of this system were, in retrospect, very obvious. The users (students in the course) had no general model of what they could use the system for, so they continually asked questions such as "Will it do X?". The same problem arose with single items on the menu (e.g., students did not understand what "calculating proportions of a population in a specified range" meant in practice). What input was required? What did the output look like? Who specified the range? And what is a range anyway? Clearly, the students needed an image of what the system could do for them in intuitive terms, so some teaching is necessary.

Similarly, the terms used by PROP to elicit information, such as object, population, and measurement, were not clear to the users. And although with a certain amount of help many of the students did succeed in obtaining answers to specific questions from PROP, they did not appreciate the importance of the assumptions underlying the answer. The answer was correct in their minds, because it was produced by the computer, but also mysterious to them, because they had no idea of the rationale behind it (see Wood, 1989 for further details).

APPENDIX 2: A RESAMPLING PROCEDURE FOR DERIVING LIMITS FOR SPC CHARTS

Resampling, or "bootstrapping," is a (computer-intensive) approach to estimating sampling error by drawing random "resamples" from a sample of data (see, Gunter, 1991; Kennedy & Schumacher, 1993; Noreen, 1989). Resampling can be used for estimating control limits for statistical process control charts.

The first step in producing a control chart for the mean is to calculate and plot the means of a series of (say) 20 samples of six measurements each on a graph. The data from all these samples (i.e., 120 measurements) is then used in the resampling process. The results shown in Figures 2 and 3 were produced by a simple computer program.

The resampling procedure works by taking random "resamples" from all the data (with replacement, otherwise it would run out of data fairly quickly). These are called "resamples" because the samples are being sampled again. The mean of each of these resamples is then calculated. For example, the first such resample in one example was 1737, 1716, 1753, 1622, 1701, and 1759. The mean of this resample was 1715. The computer program then takes a large number of further resamples from the original sample of 120 measurements. For example, the first 100,000 are shown in Figure 2. In Figure 2, the 99.8% interval goes from 1669 to 1863; the 95% interval goes from 1702 to 1816; and the median is 1753.

```
X represents 916 resamples; - represents fewer than this.
 1547 to 1575
 1576 to 1604
 1605 to 1633 -
 1634 to 1662 -
 1663 to 1691 X-
 1692 to 1720 XXXXXXXXXX-
 1721 to 1749 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-
 1750 to 1778 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-
 1779 to 1807 XXXXXXXXXXXXXXX-
 1808 to 1836 XXX-
 1837 to 1865 -
 1866 to 1894 -
 1895 to 1923 -
 1924 to 1952 -
 1953 to 1981
```

Figure 2: Means of 100,000 resamples of 6

The purpose of resampling is to see how much these means will vary if the resamples are drawn at random. It is then possible to compare the means of actual samples with this pattern. If the mean of a sample is right outside the pattern of the random resamples, a reasonable conclusion is that there is a "special cause" operating, which should be checked.

In Figure 2, 99.8% of resample means are between 1669 and 1863, so these would form the action limits on the control chart, which can be plotted in the usual way.

The same procedure can be applied to any statistic based on a random sample of data. It could be used to estimate the control limits of the range chart that would normally accompany the above mean chart. Resampling can also be used to estimate control limits for *p* charts. There are obvious advantages (e.g., increased clarity and reduced training requirements) of having one procedure that can be used in a number of different situations.

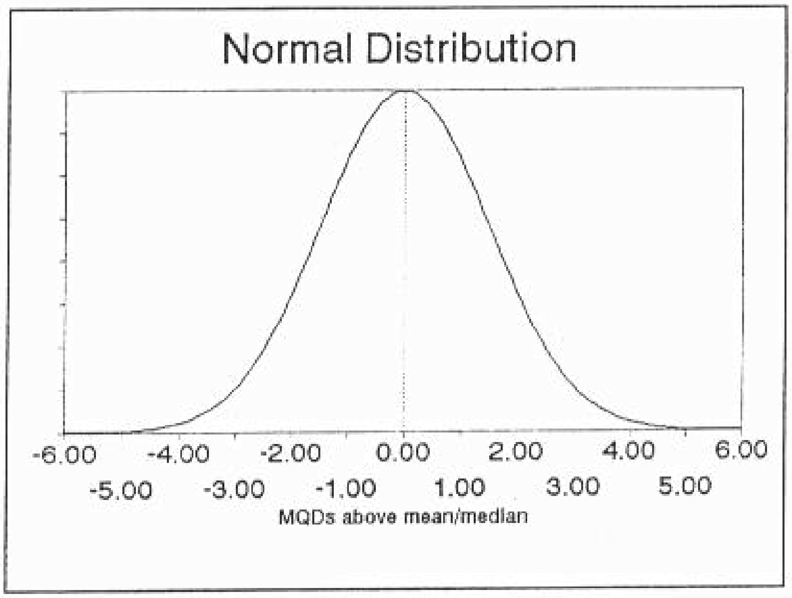
Figure 3 shows the corresponding pattern for the ranges. Notice that this does not follow the symmetrical normal shape. The action limits are at 23 and 492. In Figure 3, the 99.8% interval goes from 23 to 492; the 95% interval goes from 50 to 388; and the median is 152.

Statistical methods based on the resampling principle, such as the one described here, are gaining in popularity because they have a number of advantages (Gunter, 1991; Kennedy & Schumacher, 1993; Noreen, 1989), all of which are relevant here. Some of the advantages include that resampling methods are more transparent and do not rely on an understanding of mathematical probability theory, and that they are robust against inappropriate assumptions about population distributions because the only assumption made is typically that the sample is an adequate surrogate for the population for the purpose of estimating sampling error (see Wood, 1995 for further details).

X represents 385 resamples; - represents fewer than this.
0 to 28 -
29 to 57 XXXXXXXXXXX-
58 to 86 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-
87 to 115 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-
116 to 144 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-
145 to 173 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX-
174 to 202 XXXXXXXXXXXXXXXXXXXXXXX-
203 to 231 XXXXXXXXXXXXXXXXXXX-
232 to 260 XXXXXXXXXXXXXXXXX-
261 to 289 XXXXXXXXXXXXX-
290 to 318 XXXXXXX-
319 to 347 XXXXXXX-
348 to 376 XXXXXX-
377 to 405 XXX-
406 to 434 XX-
435 to 463 X-
464 to 492 -
493 to 521 -
522 to 550

Figure 3: Ranges of 100,000 resamples of 6

APPENDIX 3: NORMAL TABLES BASED ON QUARTILE STATISTICS



("MQD" is the median-quartile deviation: ie the semi-interquartile range.)

The table gives the proportion in each tail of the distribution. For example, it indicates that 36.79% are more than 0.5 MQDs above the median, and a similar number are less than 0.5 MQDs below the median.

MQDs from median	Proportion in tail	MQDs from median	Proportion in tail	MQDs from median	Proportion in tail
0.0	50.00%	2.2	6.90%	4.4	0.17%
0.1	47.33%	2.3	6.04%	4.5	0.14%
0.2	44.66%	2.4	5.27%	4.6	0.12%
0.3	42.00%	2.5	4.58%	4.7	0.10%
0.4	39.37%	2.6	3.96%	4.8	0.08%
0.5	36.79%	2.7	3.41%	4.9	0.07%
0.6	34.27%	2.8	2.93%	5.0	0.05%
0.7	31.83%	2.9	2.51%	5.1	0.04%
0.8	29.46%	3.0	2.14%	5.2	0.04%
0.9	27.19%	3.1	1.82%	5.3	0.03%
1.0	25.00%	3.2	1.54%	5.4	0.02%
1.1	22.92%	3.3	1.30%	5.5	0.02%
1.2	20.93%	3.4	1.10%	5.6	0.02%
1.3	19.06%	3.5	0.92%	5.7	0.01%
1.4	17.28%	3.6	0.77%	5.8	0.01%
1.5	15.62%	3.7	0.64%	5.9	0.01%
1.6	14.06%	3.8	0.54%	6.0	0.01%
1.7	12.61%	3.9	0.45%	6.1	0.01%
1.8	11.27%	4.0	0.37%	6.2	0.01%
1.9	10.03%	4.1	0.31%	6.3	0.00%
2.0	8.89%	4.2	0.25%	6.4	0.00%
2.1	7.84%	4.3	0.21%	6.5	0.00%

APPENDIX 4: A SPREADSHEET MODEL OF THE NORMAL DISTRIBUTION

Figure 4 is a screen from a Quattro Pro spreadsheet set up to plot a normal distribution and calculate the probability of a value being in a specified range. The user need only enter appropriate figures in the shaded cells, and then the probability (95.5%) will be calculated and the graph drawn. Any of the figures can be changed, and the probability will be recalculated and the graph redrawn. (The formulae are on a different part of the worksheet which the user has the option of examining.) A similar spreadsheet could be set up using quartiles as input in place of the standard deviation (see Wood, 1992 for further details).

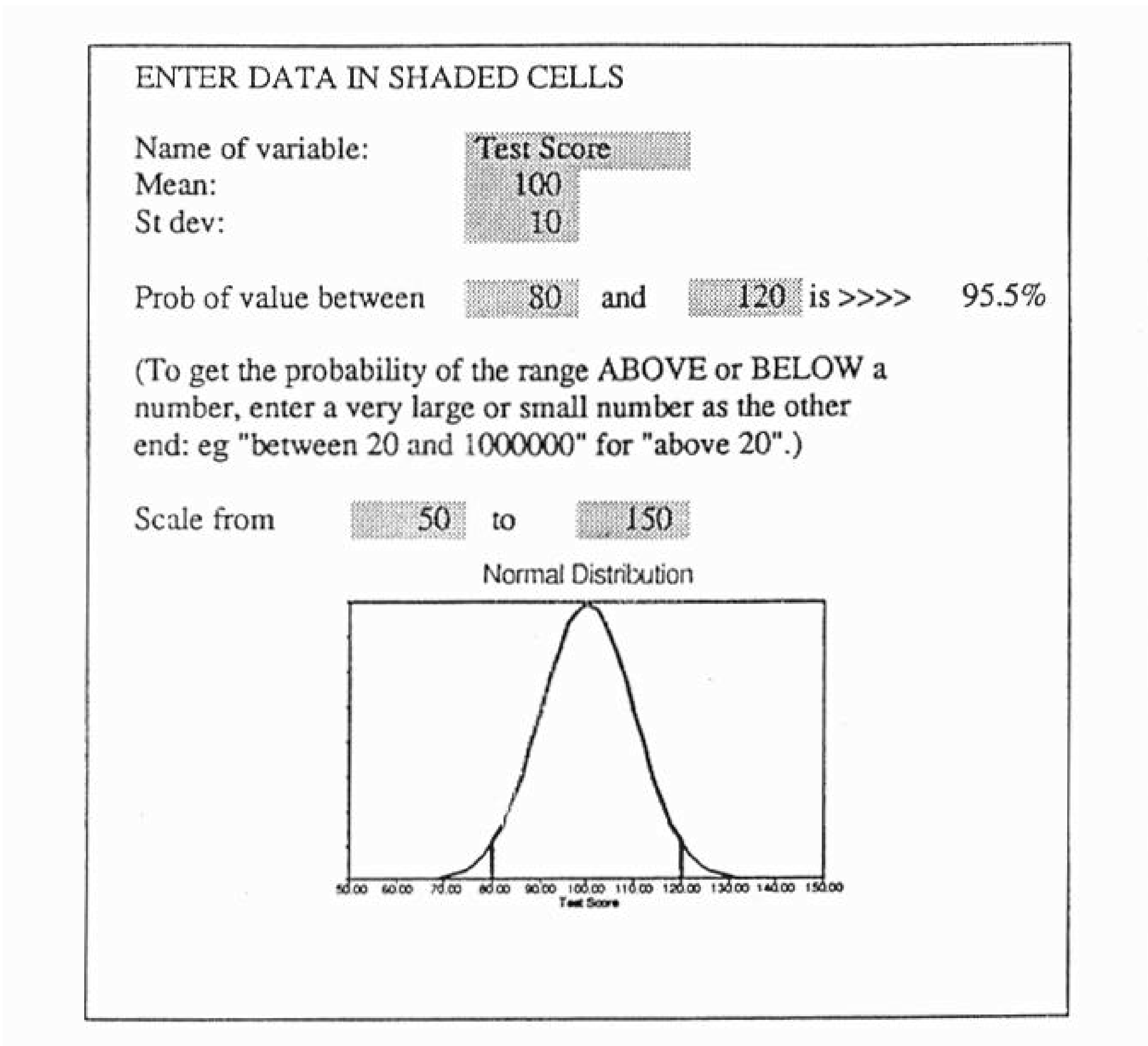


Figure 4: Spreadsheet model of normal distribution