STATISTICAL THINKING FOR EFFECTIVE MANAGEMENT ®

J A John
University of Waikato, New Zealand
D G Johnson
University of Loughborough, England

Most managers do not instinctively think statistically, mainly because they are not convinced that statistical thinking adds any value to management and decision-making. Traditional business statistics courses tend to reinforce this view by concentrating on mathematical detail and computation. Without the ability to think statistically, and to understand and interpret data, managers have to resort to gut reactions, which are invariably misguided and unreliable. In this paper we advocate a problem centred approach to teaching statistical thinking based on realistic business examples. Students must be thoroughly involved in the learning process, and encouraged to discover for themselves the meaning, importance and relevance of statistical concepts. Time should be devoted to thinking about the key issues, and for significant interaction both between student and teacher and also, more importantly, between the students themselves.

INTRODUCTION

Many managers, whether in manufacturing or processing industries, in service organisations, government, education or health, think that statistics is not relevant to their jobs. Perhaps it is because they see statistics as a collection of complicated techniques that they do not perceive as adding value to the process of management and decision-making. Traditional business statistics books and courses tend to reinforce this view by concentrating on mathematical detail and heavy computation, and largely fail to address the key issue of why managers or business students need to have any understanding of what they see to be as an irrelevant, difficult and technical subject.

However, this is not just an issue for statistics. Many students find accounting and financial concepts equally hard to understand, but there is no question in their minds that this is a necessary skill for effective management. In most situations, money is the 'bottom line', and it is not difficult to convince students that they need to thoroughly understand the language and concepts of cost control and financial management. Likewise we, as teachers of statistics, must be able to demonstrate to our students why the ability to think statistically is an essential and important skill for any successful manager.

There is a clear recognition that we need to change the way we teach business statistics. This was identified as long ago as 1986, at least in the USA, with the first of a series of annual conferences on the theme of *Making Statistics More Effective in Schools of Business.* In their overview to the first conference Easton, Roberts and Tiao (1986) note that there is "substantial dissatisfaction with much of current teaching of business statistics, especially as reflected in the poor selection of topics in popular textbooks and the limited opportunity for students to work with real data or to make serious use of statistical computing." There was also a consensus that:

1. Students are most effectively motivated by seeing statistics at work in real applications, problems, cases, and projects.
2. It is desirable to reduce emphasis on formal theory of statistics and probability and to increase emphasis on applications. Formal testing of hypotheses, in particular, receives too much attention currently.
3. Certain topics now omitted or treated lightly in basic courses such as time series, quality and productivity (including process control and experimental design), sampling, and report writing deserve consideration for an expanded role.

Unfortunately, this foresight seems to have gone largely unnoticed, and there has been little change to the way that business statistics has continued to be taught. Even now, 16 years on, the approach is still largely based on a general prescription for the teaching of statistics, whether it is to business students, engineers, scientists or social scientists. This traditional approach is clearly not relevant to the needs of managers and business students, if in fact it was ever relevant, and furthermore has a reputation for dullness and irrelevance that is exceeded only by the reality.

Insufficient attention has been paid to developing a clear understanding of why statistics is relevant and important to managers. There needs to be a shift away from the traditional course, to one that puts an emphasis on statistical thinking, and on an understanding of basic statistical ideas and concepts. In addition, learning has to be clearly outcome oriented and be focussed on managerially relevant issues, such as quality and process improvement, to which managers and business students can readily relate. Above all, the methods used must be designed to engage students' interest and involvement in a way that will create deep and lasting understanding rather than superficial recall that is immediately forgotten once the examination has been passed.

We shall first discuss what we regard to be the traditional business statistics approach, and examine why we believe it is not relevant to the needs of today’s managers. Then we shall consider a prescription for teaching a course designed to motivate statistical thinking.

TRADITIONAL BUSINESS STATISTICS APPROACH

A traditional first course in business statistics can usually be characterised by the following features.

- It places great emphasis on understanding a range of standard statistical techniques. This is often built around traditional probability theory and distributions, leading to formal hypothesis testing and then to more complex situations involving a range of multivariate techniques.
- It gives prominence to the mathematical theory of statistics, with much use made of algebraic representation and complicated formulae. Ask any student what a standard deviation is and they will probably be able to repeat (more or less correctly!) a suitable formula, but will have little or no idea what the standard deviation says about a set of data. Even if business students are well equipped (and most are not) to handle the language of mathematics, the use of complicated formulae is not relevant to their needs.
- It involves detailed calculation of quantities that may be meaningful to statisticians, but not to most managers. For many students, using statistics equates to performing such calculations. Why they are doing it, and what the results indicate, are often completely overlooked.
- It focuses on the mechanics of data analysis. Calculating a line of best fit to a set of bi-variate data is not an end in itself, but is just one way of trying to understand patterns in data. Data analysis is almost always necessary, but seldom sufficient. It is occasionally helpful to do simple calculations to get the 'feel of things', but that should not become an end in itself. What is needed is to be able to get meaningful information from data as a basis for effective management.
- It largely ignores the widespread use of computers in all business activities. Word processing, the use of spreadsheets and presentation packages are skills that most managers, and many business students, possess. Carrying out calculations ‘by hand’ still seems to be the primary approach of the traditional course.
- It tends to be illustrated by totally unrealistic, if not meaningless, examples. The examples used should be the key to motivating students’ appreciation of the importance of statistics. Simplistic calculations on data that relates to artificial examples will only tend to re-enforce students’ belief that statistics is irrelevant for realistic decision-making and problem solving.
- It includes a range of topics that are only of marginal relevance to managers. The main culprits here are probability and distribution theory and hypothesis testing. Considerable space is devoted in most business statistics books to these topics. While distribution theory forms the core of statistics, it is the domain of the statistician not the manager. Most managers will seldom feel the need to test any hypotheses, as knowing whether to accept or reject a hypothesis is not a useful basis for any effective action. In fact, there is often little to distinguish a traditional business statistics course from other introductory courses in statistics; only the examples are different.

- Classes are often totally passive, operating in 'push' mode where the lecturer talks and the students listen (or not!). The extent of the interaction is often limited to the occasional query as a result of a final "Any questions?" invitation.

Scholtes (1998) questions why people are so resistant to using data in planning and problem solving. He puts forward a number of theories. One is that statistics has been traditionally taught in such a way that the only point of interest is whether students will be bored to death or scared to death. By the time people become managers, they have taken a vow of statistical abstinence. It is imperative that we change the way we teach business statistics. What is the alternative?

TEACHING STATISTICAL THINKING

Experience working as consultants in industry and in teaching statistics to business students and practising managers for many years leads us to suggest a number of general principles in designing a course on statistical thinking. It is important that students think about *what* needs to be investigated, and *why*. This is best done by means of simple, yet realistic, management scenarios that give rise to the issues involved in understanding and interpreting data. Students should be encouraged to decide for themselves what is needed, rather than be told what to do.

One of the key issues that managers should focus on is variability. Managers need to understand the nature of variability; otherwise they end up making wrong or inappropriate decisions that can be very costly to their organisations. Students have to be aware of the consequences of failing to appreciate of the importance of variability, and of how to manage it. This understanding is fundamental to effective management and decision-making. Certain statistical concepts, mainly to do with the measurement and interpretation of variability and uncertainty, must be clearly understood.

We should not get side-tracked into extensive data analysis and hand calculation merely for the sake of it. Instead, students should be encouraged to use appropriate computer-based tools for data analysis. Spreadsheets are ideally suited for this purpose. A package such as Excel can be used to great effect for most of the simple data analysis that is required in an introductory business statistics course. Furthermore, many students will be familiar with the main features of a spreadsheet, such as Excel, and will only need to be shown some of the more advanced features that are used for data analysis.

A course that focuses on statistical thinking would need to be application oriented, and based around topics such as variation, problem solving, presenting and modelling data, sampling and estimation, attribute and related data, forecasting, statistical process control and process improvement. It is important to keep in mind that students must be continually motivated to see the need to examine data in various ways, and to have a desire to solve relevant problems. Any theory should be introduced only if motivated by a particular problem or by the need to understand a particular concept. For example, a thorough introduction to the normal distribution is necessary because of its pivotal role in modelling data and understanding sampling error. It should not, however, be taught as an end in itself. Likewise, key ideas from the binomial and Poisson distributions are a necessary precursor to understanding variation in attribute data. The approximate normality of these distributions, along with their standard deviations, is an important concept. Being able to use the binomial and Poisson formulae to calculate probabilities, however, is of much less importance to managers.

Explaining the variation in a set of data is a fundamental activity of statistics. It is what we try to do in regression analysis or with analysis of variance methods. With attribute data, which arises frequently in a business context from such sources as market surveys, consumer research and product testing, again we attempt to explain the variability in the data. Of course, we may ask whether a particular component or factor is explaining a significant amount of variability, and for this we formally use the F and chi-square tests for examining possible relationships between interval and attribute variables. Experience suggests that the notion of significance is a difficult concept for students to grasp, but one that is necessary if they are to draw sensible conclusions from their data. What is important, however, is understanding the meaning of a significance probability rather than computing chi-square or F values. For most students, this requires careful explanation, and much discussion of meaningful examples, if they are to correctly interpret the information contained in a set of data.

SOME LEARNING CONCEPTS

In teaching courses on statistical thinking over many years we have found three learning concepts particularly important and useful in helping business students to appreciate the importance and value of statistics. They are *interaction, discovery* and *experimentation*.

Teaching statistical thinking requires time to think, and to discuss statistical ideas and concepts. An approach that creates an impression of "here is a problem and this is what you do to solve it" is unlikely to succeed. A course in statistical thinking must involve significant *interaction* both between student and teacher and also, more importantly, between the students themselves. Students will learn statistical principles more effectively, just like any other subject, by active involvement in the learning process. Interaction in a small class of managers or MBA students is relatively easy. In particular, these students have experience that they can draw on and use to respond to the given challenges. However, even in a large class of, say, 250 first year business students we have found it possible to have some useful interaction with the class. While teamwork is difficult, students can discuss issues with those sitting next to them or in their immediate vicinity. We have found that students enjoy such discussions, as the following comments from some students in a recent class show.

- I think that interaction is good. It gets us involved, which makes us work and think more.
- I like interaction because it allows me to see whether my thoughts are the same as my peers or not.
- Having time to discuss questions and ideas at the time really helped my understanding.
- Interaction between the lecturer and student allows greater collaboration of ideas and concepts.
- You can discuss things with friends and get a better understanding of how a particular problem works, rather than nut it out by yourself.
- It also encourages students to do some thinking as well as listening.

Discovery for oneself is another key principle. For instance, the concept of measuring variability is fundamental to data analysis, and the notion of the standard deviation is a critical one. In many traditional statistics books, the approach is essentially "we need to measure variability, and here is a formula that does it". This is unlikely to be successful as few students will be able to make the connection between an abstract algebraic formula and something that measures variability in data. Is the formula strictly necessary, or is it more likely to confuse than to enlighten? Almost any calculator or spreadsheet will give the standard deviation of any set of numbers at the press of a key or the selection of the appropriate function. The time will be much more profitably spent getting students to discover for themselves how they would measure variability. One way of doing this is to present the data in the form of a run chart, rather than the more usual histogram, so that the idea of a range easily emerges, and this in turn leads naturally to looking at deviations (or residuals), giving rise to the mean deviation and ultimately the standard deviation. As another example it is interesting to challenge students to predict the value of some response variable given a value of an explanatory variable. Given a scatterplot of the data most students will use a straight edge to draw a line that “fits” the data. This type of exercise readily leads on to least squares and the assumptions that need to be made.

Many of the fundamental concepts in statistics can be usefully illustrated by appropriate *experiments, simulations* or *demonstrations*. Probably the most useful and widely used experiment is the well-known "red beads" experiment devised many years ago by Dr W Edwards Deming (1986). This is a good way to start off a course on statistical thinking as it presents underlying ideas about process variation in a very simple and entertaining way. It can be run in a one hour class, with any size group, and involves the students in generating data by sampling a population consisting of a large number of beads, some of which are red (defective) and the majority white (good). A number of students (typically about 5) take it in turns to "produce" a set of 50 beads over a number of days (again usually 5), thereby generating about 25 sample results in total. The production process consists of dipping a wooden or plastic "paddle" into a thoroughly mixed box of beads and collecting a sample of 50 beads in a series of circular depressions set into the paddle in 10 rows of 5. The results produced by each worker (student) are counted, checked and recorded by other students involved in the experiment. As the population consists of many hundreds of beads, the situation should in theory conform closely to a binomial distribution. Furthermore, because the sample size (50) is large, the distribution should be approximately normal. However, it has been observed in practice that the results do not necessarily conform to expectation. For example, one particular set of beads consisting of 1200 white and 400 red beads, which should in theory produce an average of 12.5 red beads each time, consistently averages about 10. The reason for this bias is not clear, but it certainly provides an interesting and important discussion point. Similarly, the students are invited to think about and discuss whether the composition of beads in the population has any bearing on the variation in the results produced. It is quite surprising to note that even the non-mathematically inclined students realise, without too much prompting, that as the proportion of red beads increases, so does the variability. It is also possible to draw out quite easily that the greatest variability arises with a 50:50 composition of red and white beads; a result which is important when considering the margin of error in a sample proportion.

Overall, the red beads experiment is a fun way to start any course on statistics, which sets the tone for the course and provides an example of numerous important concepts. It can be continually referred to throughout the course when discussing topics such as normal (and binomial) distributions, margin of error, process control and process improvement. Although probably the most important experiment, the red beads experiment is just one of many that can be used. We use other experiments to illustrate the importance of using random sampling rather than judgment sampling, the meaning of a margin of error, least squares and an understanding of R^2, the benefits of reducing variation, the effects of tampering with a process, and the use of designed experiments to improve a process. There are a number of published sources of experiments and class-based activities that can be used to illustrate key learning points. A good starting point is the book by Scheaffer et al. (1996).

It is generally better, wherever possible, to run the actual experiment in class. This gives the benefits of active participation in the learning process, and also allows students to see things 'for real'. Unfortunately this is not always possible, either because the particular equipment required for the experiment might not be available or because a student might be studying alone. In many cases, however, a good alternative to the actual experiment is to use a computer-based simulation, which reproduces as nearly as possible the actual experimental process. It also allows typical experimental results to be generated so that the outcomes of the experiment can be seen. With a little ingenuity, and some facility with Excel and Visual Basic, a range of computer-based experiments and demonstrations can be developed to simulate the real experiments mentioned above.

Where students have access to web-based learning resources, the Rice Virtual Statistics Lab (RVLS) at http://www.ruf.rice.edu/~lane/rvls.html offers a comprehensive statistics learning support package including numerous practical demonstrations of key statistics concepts as well as a range of data sets and case studies.

CONCLUSIONS

There are a number of important elements to consider when developing a course in statistical thinking. Allow students time in class to discuss and explore issues, problems and alternatives. Material should be presented in the context of typical managerial situations, rather than as a sequence of theoretical topics that are then 'applied' to 'problems'. Relevant and meaningful examples should be used to stimulate thought and to illustrate key concepts. Encourage students to *talk* both to each other and to you, the teacher. Avoid the language of mathematics. It is our language, not theirs! Avoid unnecessary technical detail wherever possible. Only give formulae when absolutely necessary. If calculations can be done on the computer, using Excel for instance, do them that way rather than relying on manual methods. Use whatever resources are available, or that you can develop, to make learning more interesting and long lasting. Students will remember taking part in the red beads experiment more readily than they will recall a lecture on variation!

A recent text by John, Whitaker and Johnson (2001) has attempted to capture these essential ideas. It is based on the authors' extensive experience in teaching statistics to managers and business students, and has been developed over 7 years with students in New Zealand and the UK. It is written in the form of a workbook, which students use interactively in class to write down their thoughts, ideas and answers to the large number of questions designed to stimulate their learning and test their understanding. In addition, the book is supported by an instructor's manual, PowerPoint slides presenting the key learning points, and a set of Excel files incorporating both sample data sets and various computer-based simulations and experiments. The book has been designed as a first introduction to statistical thinking, and has been used successfully not only with first year undergraduate students, but also MBA students and practising managers.

REFERENCES

Deming W.E. (1986). *Out of the crisis*. Cambridge: Massachusetts Institute of Technology Center for Advanced Engineering Study.

Easton G., Roberts H.V., & Tiao G.C. (1986). *Overview and summary*. Proceedings of Making Statistics More Effective In Schools Of Business. University of Chicago.

John J.A., Whitaker D., & Johnson D.G. (2001). *Statistical thinking for managers*. London: Chapman & Hall/CRC.

Scheaffer, R.L., Witmer, J., Watkins, A., & Gnanadesikan, M. (1996). *Activity based statistics*. New York: Springer.

Scholtes, P.R. (1998). *The leader's handbook*. New York: McGraw-Hill.