Teaching Statistical Concepts, Fundamentals and Modelling

KEYWORDS:
Teaching;
Statistical literacy;
Course design.

Timothy E. O'Brien
Loyola University Chicago, USA
e-mail: tobrie1@luc.edu

Summary
This article considers some issues in designing a course focusing on statistical concepts rather than memorizing formulae.

◆ INTRODUCTION ◆

Some time ago, I was assisting a genetics colleague with the analysis of his fruit fly data collected over several locations in Australia. The response variable in his study was binary (success or failure) and the explanatory variable was degrees of south latitude. After the data were obtained, the colleague claimed he could easily 'crunch the numbers' himself, and then proceeded to produce and analyse his derived 2×2 contingency table in which one row corresponded to 'low latitudes' and the other to 'high latitudes.' He was elated when his analysis produced 'statistical significance', by way of the all-important '$p < 5\%$'. It was then that I started to appreciate the great disservice that we statistics instructors may engage in when we teach 'number crunching' statistics courses wherein formulae are emphasized over concepts, wherein rote memorization takes precedence over reasoning.

After several years of teaching introductory statistics courses, and reflecting on the larger issue of statistical literacy in society, I have shifted my focus over the past several years away from the '*whether*' and '*why*' to just '*how*' one can achieve that. Thus I have come to realize that every educated citizen does indeed need to have a basic level of statistical literacy and have started to explore how specifically we can help realize or facilitate this goal. Clearly, when introductory courses are taught emphasizing only the successful use of a given formula (or a set of steps), and failing to underscore the underlying concept and statistical model, these courses perpetuate the misconception that the key to successfully analysing a set of data amounts to using *the* 'correct formula'.

My genetics colleague no doubt had taken a statistical methods course similar to the one I taught years ago to science and social science graduate students using Ott (1993) and akin to the introductory biostatistics course I have taught to biology undergraduate students using Samuels and Witmer (2003). These texts are typical of introductory texts in that they cover the usual topics, including one- and two-sample paired and independent t tests, categorical methods, nonparametric methods, linear regression and ANOVA. From the student's perspective, through exams and homework these courses do *de facto* emphasize choosing the proper statistical tool from a statistical toolkit and successfully using that tool. Of course, these courses also underscore the critical assessment of any necessary assumptions and require that conclusions be conveyed in clear nontechnical terms, but the focus is much more on methods than on concepts. I have also had the good fortune to teach basic courses in statistical methods in countries other than my own (most recently in Thailand), and these experiences have further impressed on me the need to revise and rethink our goals in statistics courses so as to place more of an emphasis on conceptualization and understanding. In this note, I discuss concrete ways of incorporating statistical literacy into our courses.

One tool that I have used to stretch students beyond the usual teaching paradigm encountered by my genetics colleague in a statistical methods course is to have students write a course paper in which they discuss and critique the statistical techniques used in research articles of their own choosing. This project, discussed more thoroughly in O'Brien (2005), serves to help students see the usefulness of statistical methods in practice, to challenge the belief that a given data set can be analysed using only one statistical technique, and to observe the misapplication of statistical techniques in the scientific literature. Invariably, students take away from this exercise a keen appreciation of 'best' or 'powerful' statistical methods; many then go on to take additional courses in applied statistics. Further evidence of this trend is noted in the marked increase in enrollments in our new 'minor' programmes in applied statistics – a 'minor' is awarded to a Loyola student who completes four statistics courses.

In contrast to these biology and science students – students who typically go on to become 'producers' of statistics – I also recently had the opportunity to teach an introductory statistics course to humanities and social science majors (i.e. 'consumers' of statistics); throughout the remainder of this paper we focus on this specific course. More precisely, it was my goal in this statistical literacy course to depart from any emphasis on memorization and the successful implementation of a given statistical technique (such as those listed above), and to highlight and underscore concepts and understanding associated with these techniques. This was accomplished using the writing assignments, group projects, hands-on and computer activities discussed below.

TOPICS COVERED IN THE STATISTICAL LITERACY COURSE AND ASSESSMENT

When I first joined the faculty at Loyola University Chicago in 1998, the textbook in use for the statistical fundamentals course was an earlier edition of Freedman et al. (2007). For those unfamiliar with this text, it is indeed novel in that it does an excellent job of downplaying the use of involved formulae and calculations, emphasizing underlying statistical concepts instead. Thus for example, instead of conveying the idea of a sample standard deviation with the formula

$$\sqrt{\frac{\sum(y_k - \bar{y})^2}{n-1}},$$

students are asked to find it using the three-step process: (1) find the sample mean and the deviations for each score, (2) square the deviations and average these squares using $(n - 1)$ in the denominator and (3) take the square root of this average. Students can easily remember this root-mean-square (RMS) process, and this concept is encountered later in the course in the chapter on regression. Some educators feel that all college students should be able to calculate a standard deviation and others do not. In my own experience, I get a much higher proportion of correct answers for the nonscience students considered here when the RMS strategy is used instead of the above formula.

Equally noteworthy, the Freedman et al.'s (2007) text (and approach to learning statistical concepts) also introduces so-called *box models* as a metaphor for the population under study. For example, in the chapter that discusses Mendel's genetics work with pea plants involving crossing two heterozygous parents (each with one dominant trait, denoted A, and one recessive trait, denoted a), the box model for the phenotype of the next generation contains one a and three As, thereby conveying the probability of 75% of observing a dominant offspring. Box models are also used in Buntinas and Funk (2005) and lay the foundation for statistical modelling; see also Moore and Notz (2006).

Since I found that many of the examples in Freedman et al.'s (2007) text are somewhat dated, as it lacks extensive resources for instructors such as applets and Microsoft PowerPoint presentations and since I wanted to further de-emphasize calculations, I instead used the text by Utts (2005a) for this course. This latter text focuses on statistical concepts and is coupled with excellent detailed resources for instructors and with an activities manual, Utts (2005b), which I found very helpful as an aid to facilitate learning. I also augmented this text with other teaching materials discussed below.

In the class, we spent roughly two-thirds of the 50-minute class meetings covering the text material (approximately one chapter per class), and the remaining class periods involved in in-class exercises or in the computer laboratory. Final grades were calculated using the following breakdown. I based the student's grade on a midterm exam (22.5%), final exam (22.5%), homework and quizzes (20%), group-work and participation (10%), 'mini-projects' taken from the text (15%) and a final course project (10%). Exams were made up of conceptual questions (short answer, multiple choice and fill-in-the-blank problems) and exercises. Calculators were required for homework and exams (although they were used much less than in a typical 'methods course'). As this was my first experience teaching such a purely conceptual statistics course, I followed the text and suggested class activities quite closely. We covered the topics of (1) obtaining reliable data (discussing experiments versus observational studies and bias), (2) representing data graphically (covering regression and 2×2 contingency tables), (3) probability basics and (4) estimation and hypothesis testing (including finding and interpreting confidence intervals).

I felt that the text did a good job of sharpening students' eyes to spotting potential biases in news and other studies and stories, paved the way for hypothesis testing by introducing the basics of testing for 2×2 contingency tables, and did not overdo things with details by just covering the basics of probability. Although the text by Utts (2005a) was the primary textbook for the course, I augmented this material by introducing *box models* (e.g. to illustrate the Central Limit Theorem for independent Bernoulli trials), by assigning outside readings from Huff (1993), Cobb and Gehlbach (2006), and Peck et al. (2006), by using the Minitab statistical package when needed and engaging in computer activities to illustrate the 'German Tank Exercise' discussed in Bullard (2003), and by having students gather and interpret their own data that they found of interest (for their course projects).

A colleague and I each taught a section of this concepts course to approximately 25 students, and the students fared quite well in terms of grades. This is not surprising since Loyola students tend to be quite bright and well trained in writing, accustomed to active and group learning, and since the course emphasized concepts over memorization of mathematical methods. Although our results may be merely anecdotal, student evaluations of this statistical literacy course were surprisingly very high and quite positive. To better give an idea of the course we offered, following are some further details regarding activities, teaching methods, writing assignments and projects.

◆ CLASSROOM ACTIVITIES AND OTHER TEACHING AND LEARNING TOOLS ◆

As an instructor trained in mathematics and statistics, and with only the basics of teaching pedagogy given in my Peace Corps training in West Africa, I found the classroom activities given in Utts (2005b) tremendously beneficial by providing me with the means to underscore the ideas and techniques discussed in the classroom lecture and giving students the opportunity to learn from one another. For example, students learned about the strategy of staying with one's original choice or switching to another choice in the 'Let's Make a Deal' activity (modelled after a television game show popular in the USA and Canada and used to demonstrate basic probability). In another instance, confidence intervals were illustrated via an in-class activity wherein each student dropped a tack 100 times and counts were made of the number of times the prong landed facing up (deemed a 'success' for this exercise). This exercise also illustrated how confidence intervals became narrower when the results are pooled over the four students in each group, that is, when the sample size increases from 100 to 400. In several activity sessions, students were asked to read through and discuss abstracts of research articles (or the actual articles themselves) highlighting potential biases, confounding and interacting variables and so on. At the end of these sessions, students would often note how many of these potential problems were ignored by the authors and thus how questionable were the findings given in these studies and reported in a newspaper or on the Internet.

In introductory courses, educators also invariably encounter students falling victim to what Utts calls the 'confusion of the inverse'. On the first day of class, I had the students answer the question 'Suppose 1 in 1000 people have a disease. A test for it has a 10% false-positive rate and a 10% false-negative rate. If someone tests positive for the disease, the chances that they actually have it are about what percentage?' The roughly three-quarters of the students who answered '90%' had thus fallen prey to this confusion. Interestingly, the approach advocated in Utts (2005a) to convey the framework to correctly answer this question – as well as to introduce students to conditional probabilities and Bayes' Rule in general – is to use a specific 2×2 table. As a quantitatively minded individual, I had always used probability trees to get these concepts across. Again, based only on my anecdotal experiences, I was surprised to note that roughly half of my students favoured this 2×2 approach whereas the other half found the more traditional use of probability trees more beneficial. It is interesting to note that even for such a homogeneous (non-mathematical) group of students, what worked to convey an idea for one student did not carry over to another student.

◆ WRITING ASSIGNMENTS AND PROJECTS ◆

Since my audience was humanities and social science majors, they really did not mind that 25% of their final grade was based on assignments involving writing and that this course was writing intensive: they felt that the course played to their strengths. The text by Utts (2005a) has mini-projects at the end of each chapter, and at three instances during the semester students completed one of these related to part of the text. For example, one of these asked students to find a research article discussing an observational study and to write a one- to two-page summary of the study and to mention potential biases and the possible ramifications of these. As an additional assignment, I had students complete the second assignment given in Jordan (2004) in which students are asked to write a nontechnical letter to a friend explaining and interpreting in layman's terms the *p*-value associated with a randomized placebo-controlled study and to discuss its importance. It is precisely an ability such as this one that goes a long way to distinguish a statistically literate citizen from an illiterate one; I was very pleased to see this skill developed in my students in this course. Furthermore, the final course project/paper required that students obtain their own data on a quantitative variable for each of two groups, summarize the data, test hypotheses and write up their findings.

◆ CONCLUSION ◆

Even though there may be some truth in the statement that 'all [statistical] models are wrong but some are useful' (Box 1979), a course in statistical methods that ignores statistical modelling and that fails to underscore the underlying concepts does a great disservice to our students; these courses are seen as a jumble of formulae that leave students with the impression that statistics is merely the 'number crunching' my genetics colleagues had in mind. On the contrary, the real strength of statistical science is its ability to develop new methods to directly aid researchers and decision-makers, and as educators we need to emphasize this flexibility and adaptability in our courses and discussions.

By way of an epilogue, encouraged by my successes in this course and some of my students, I have subsequently taught another course on statistical concepts, this one entitled *Statistics and Medical Ethics*, using the texts by Angell (2004), Avorn (2004) and Crossen (1996). The course met with good success, and will be offered again in next semester under the title *Numbers and Ethics*, and using the additional reference by Jackson and Jamieson (2007). Indeed the popularity of these courses and subjects attests to the need for a truly statistically literate populace.

Acknowledgements

I am very indebted to my colleague, Dr Kathryn Nyman, for her insights and discussions related to the statistical literacy course during the Fall 2005 semester. I also gratefully acknowledge the kind hospitality and support of the Department of Statistics, Faculty of Science, Chiang Mai University, which hosted me during my Fulbright Scholar visit in 2006–2007, as well as the generous financial support of the US Department of State, the J. William Fulbright Foreign Scholarship Program, and the Thailand – US Educational Foundation.

References

Angell, M. (2004). *The Truth About the Drug Companies*. New York: Random House.

Avorn, J. (2004). *Powerful Medicines: The Benefits, Risks, and Costs of Prescription Drugs*. New York: Knopf.

Box, G.E.P. (1979). Robustness in the strategy of scientific model building. In: R.L. Launer and G.N. Wilkinson (eds) *Robustness in Statistics*. New York: Academic Press.

Bullard, F. (2003). *The German Tank Problem*. downloaded from the Beyond the Formula archives URL: http://web.monroecc.edu/manila/webfiles/beyond/2003S022S071Bullard.pdf.

Buntinas, M. and Funk, G.M. (2005). *Statistics for the Sciences*. Belmont, CA: Thomson.

Cobb, G. and Gehlbach, S. (2006). Statistics in the courtroom: United States v. Kristen Gilbert. In: R. Peck, G. Casella, G. Cobb, R. Hoerl, D. Nolan, R. Starbuck and H. Stern (eds) *Statistics: A Guide to the Unknown* (4th edn), pp. 3–18. Belmont, CA: Thomson.

Crossen, C. (1996). *Tainted Truth: The Manipulation of Fact in America*. New York: Simon & Schuster.

Freedman, D., Pisani, R. and Purves, R. (2007). *Statistics* (4th edn). New York: Norton.

Huff, D. (1993). *How to Lie with Statistics*. New York: W.W. Norton & Company.

Jackson, B. and Jamieson, K.H. (2007). *UnSpun: Finding Facts in a World of Disinformation*. New York: Random House.
Jordan, J. (2004). *Tales from the Frontline: A Pilot Study of Writing Assignments*. Handout at 'Innovative Ideas for the Statistics Classroom' session at the Joint Statistical Meetings, Toronto, CA.
Moore, D.S. and Notz, W.I. (2006). *Statistics: Concepts and Controversies* (6th edn). New York: W.H. Freeman and Co.
O'Brien, T.E. 2005. The importance of projects in applied statistics courses. In R.J. Maher (ed.) *Innovative Methods in Undergraduate Courses following Calculus*, MAA Notes Series #67, pp. 115–125. Washington, DC: The Mathematical Association of America.
Ott, R.L. (1993). *An Introduction to Statistical Methods and Data Analysis* (4th edn). Pacific Grove, CA: Duxbury.
Peck, R. Casella, G., Cobb, G., Hoerl, R., Nolan, D., Starbuck, R. and Stern, H. ed. (2006). *Statistics: A Guide to the Unknown* (4th edn), pp. 3–18. Belmont, CA: Thomson.
Samuels, M.L. and Witmer, J.A. (2003). *Statistics for the Life Sciences* (3rd edn). Upper Saddle River, NJ: Pearson Education (Prentice-Hall).
Utts, J.M. (2005a). *Seeing Through Statistics* (3rd edn). Belmont, CA: Thomson.
Utts, J.M. (2005b). *Activities Manual for Seeing Through Statistics* (3rd edn). Belmont, CA: Thomson.