How to Make the Learning of Statistics Interesting, Fun and Personally Relevant: Using Progressive Material as Examples for In-class Analysis and to Raise Social Awareness

Kai-Lit Phua

Introduction

One of the greatest challenges in statistics education is the teaching of introductory statistics to non-specialists (Yilmaz 1996). These include undergraduate non-majors. Non-majors usually take a course in introductory statistics not because of personal interest in or curiosity about the subject but because of undergraduate degree requirements. Many non-majors feel that statistics is uninteresting, dry, not relevant to their lives etc. because of the following reasons:

- some may find statistical theory to be too abstract and technical in nature
- some may be intimidated by the use of mathematical notation in statistics textbooks (Higgins 1999)
- there are many new concepts and strange-sounding terms to be learned
- they fail to see any connection between statistical knowledge and their majors (for those who are not interested in doing quantitative research)
- some may regard a course in introductory statistics to be a bothersome degree requirement and thus approach the subject with a negative attitude (much like the foreign language requirement)

The Even Greater Challenge of Teaching Statistics to Special Kinds of Non-Majors (Such as Medical Students)

All the above reasons make the by no means easy job of introducing a technical subject to non-majors even more difficult. The challenges arising from the teaching of statistics to special kinds of non-majors are even greater (Carter 1987; Hacker 1987). I have been teaching introductory statistics to medical and pharmacy students in a medical school in Malaysia for eight years. My educational mission is a challenging one because the medical and pharmacy students are recent high school graduates (the Malaysian higher education system is heavily influenced by the British system and students are allowed to study for a medical or pharmacy degree directly after high school) from the "science stream" with little or no exposure to the social sciences. One major challenge I face is the narrow mind set of medical students, i.e. they are interested in learning about disease, pathology, clinical examination and diagnosis, treatment and management of disease etc. They are less interested in learning about public health approaches such as health maintenance and prevention of disease and disability. Very few are interested in learning about what they consider to be irrelevant subjects like statistics.

In my experience, the teaching of descriptive statistics is relatively easy. The teaching of correlation and simple regression is still manageable. However, significant numbers of students are confused by inferential statistics, i.e. hypothesis testing and its associated concepts such as Type 1 and Type 2 errors, 95% and 99% confidence intervals, degree of freedom, critical value, p-value etc. As mentioned earlier, unless they are taught well, many students find statistics to be uninteresting, dry and irrelevant to their quest to acquire a medical degree and eventually practice as doctors (Boring and Nutter 1984). Few are interested in becoming biomedical researchers who will need to use statistical techniques or medical school academicians who will need to have at least a basic understanding of statistics.

Teaching Descriptive Statistics to Students (and Raising Their Social Awareness at the Same Time)

When I teach students about the level of measurement, they are always very interested when I discuss how "race" (ethnicity) and gender are socially constructed nominal/categorical data. For example, I mention that people are always classified as "black" in the United States if they have one parent who is black and one parent who is white (regardless of whether the white parent is the father or the mother). One cannot choose to be classified as in-between, i.e. not until recent challenges by certain individuals who want to be classified as "mixed race". But in South Africa, there are "blacks", "whites" as well as "Coloureds".

Similarly, while it is customary to classify people as either "male" or "female", I mention that there are actually people who are in-between by being trans-sexual (and supposedly suffer from "gender identity disorder") or who are born with ambiguous genitalia.

When I discuss the normal distribution and skewed distributions, I ask students whether income is normally distributed and if they do not think so, whether the distribution of income is skewed to the right or skewed to the left. I also mention that governments can manipulate the official definition of the poverty line to make it appear that fewer citizens are living in poverty.

Using "In-Class Examples" and "Data Collected On-the-Spot" to Teach Inferential Statistics (and Get Students Thinking about Social Issues)

In my effort to make inferential statistics and hypothesis testing more interesting and relevant to students, I have come up with certain techniques for doing so. Basically, I use "in-class examples" and "data collected-on-the spot" to illustrate the utility of basic statistical tests such as the x^2 test of association. I "collect" data on-the-spot from the students physically present in the lecture

theatre by a show of hands and use this data to test hypotheses that are interesting and also relevant to the students.

Example:

Bearing in mind that the medical students (or pharmacy students) in my audience are predominantly teenagers, I collect data from them to test the following set of hypotheses:

H_0: There is no association between gender and satisfaction with weight
H_1: There is an association between gender and satisfaction with weight

Rationale: The research literature shows that females in general and young females in particular tend to be more likely to be dissatisfied with their bodies than males because of socialisation and mainstream media images. Females are more likely than males to consider themselves to be "overweight" or to be unhappy with their figures (negative body image). Beside these, cases of anorexia nervosa or bulimia nervosa also tend to be mostly female. Hypothesis testing using this set of statements thus becomes more interesting and relevant to many students in the audience (Smith 1998).

In a typical class of 160 medical students, about half the students are males and half are females. First, the 80 female students are asked to answer "Yes" or "No" (by a show of hands) to the question "Are you satisfied with your weight?" I use the show of hands to fill in the relevant cell in the 2 X 2 contingency table. Subsequently, the same process is carried out for the 80 male students.

The data is then analysed using the statistics software which has been loaded onto the hard disk of my notebook computer. The entire process is projected onto the screen of the lecture theatre.

If the statistical analysis shows that there is an association between gender and satisfaction with weight, I follow up by commenting that the existing research literature shows that females tend to be more dissatisfied with their weight and that this can give rise to behavioural problems and eating disorders. I

emphasise the negative influence of socialisation and mass media images on how women perceive their bodies and themselves.

Teaching More Advanced Statistics to Undergraduate Non-Majors

I also teach medical and pharmacy students more advanced statistics (later in their medical curriculum) such as relative risk analysis and odds ratio analysis and how to use 95% and 99% confidence intervals to determine statistical significance for these. The example I use very successfully at the beginning of the lecture to capture their attention is the well-known U.S. prospective study of a large number of women that revealed how dangerous hormone replacement therapy (HRT) actually is to the health of women (Rossouw et al. 2002). The statistical technique used by the researchers is actually hazards ratio analysis with 95% confidence intervals but I explain that the interpretation of the hazards ratio is similar to the interpretation of relative risk. The HRT example is used to make them aware of the dubious marketing efforts of the large pharmaceutical companies, i.e. using marketing techniques to make exaggerated claims for their drugs and their efforts to promote "off label" use of medical drugs in order to maximize profits. (Health Warning: caution should be exercised in the interpretation of research studies dealing with health risks. One would need an indication of the actual risks involved, e.g. if the baseline risk is low, doubling would not increase the risk very much. However, if the baseline risk is high, even an increase of 20% would increase the risk considerably. Furthermore, the context needs to be taken into account in order to decide if the risk is "clinically significant").

Conclusion

Although I am a medical sociologist and not actually a statistician by training, my career in the teaching of statistics has been very successful (as judged by student evaluations of my teaching). I attribute this success to the great effort I make in simplifying the topics as much as possible and above all, by making use of relevant examples to get the students interested and also to get them thinking critically about various social issues. I am confident that

this approach can be adopted with a high probability of success by other teachers of statistics who are progressive-minded.

Acknowledgement: Thanks are due to the editors for their comments. They have helped to improve the content of this article.

References

Boring, J.R., and Nutter, D.O. (1984) Analytic thinking: educating students for the practice of modern medicine. Journal of Medical Education. 59: 875-880.

Carter, R.A. (1987) Biostatistics Instruction at Meharry Medical College. The American Statistician. 41: 268-269.

Hacker, C.S. (1987) Medical statistics at the medical school of the University of Texas Health Science Center at Houston. The American Statistician. 41: 266-267.

Higgins, J.J. (1999) Nonmathematical statistics: a new direction for the undergraduate discipline. The American Statistician. 53: 1-6.

Rossouw, J.E. et al. (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA. 288: 321-33.

Smith, G. (1998) Learning statistics by doing statistics. Journal of Statistics Education. [Online]. 6(3).
http://www.amstat.org/publications/jse/v6n3/smith.html
Yilmaz, M.R. (1996) The challenge of teaching statistics to non-specialists. Journal of Statistics Education. [Online]. 4(1).
http://www.amstat.org/publications/jse/v4n1/yilmaz.html

Kai-Lit Phua, PhD FLMI
School of Medicine and Health Sciences
Monash University Malaysia
Jalan Kolej, Bandar Sunway
46150 Petaling Jaya, Selangor
phuakl@hotmail.com