USING EXAMPLES RELATED TO HIV/AIDS TO ENHANCE UNDERSTANDING OF STATISTICAL THEORY AND OF SOCIAL IMPLICATIONS OF STATISTICS

Jacky Galpin
University of the Witwatersrand, South Africa
jacky@galpin.co.za

Without practical applications of theory, students often do not understand why they are studying the techniques and what statistics is about. Service course students are often antagonistic to and afraid of the statistics courses. Some progress can be made by noting that anyone who crosses a road and survive, is an expert in probability. Relating techniques to current social issues can further assist in breaking down barriers, increasing awareness of the usefulness of statistics and enhancing academic citizenship. Specifically using HIV/AIDS examples, in a country where this is a major problem, helps students to understand the issues and how to deal with the risk. Since many are the first from their home areas to attend university, they have some standing in the community, and communicate with others. This paper outlines some of the issues and methods and examples used in both major stream and service courses by the author and by colleagues.

BASIC TENANT

The main aim of teaching statistics is to empower students to become statistically literate, and to apply evidence-based methods of reasoning and decision-making. If we don't achieve it, we have failed in our teaching.

WHY IS THIS SO IMPORTANT NOT JUST FOR STATISTICS TEACHERS?

The importance of achieving this is becoming ever greater in today's society, in South Africa at least. The government is moving towards evidence based governance, as was emphasized by Minister Trevor Manuel, Minister of Finance, at his opening address at ICOTS 6. It can be summarized by his appeal, "Don't tell me what I want to know, tell me what I need to know." At a meeting of heads of statistics offices in the Southern Africa Development Countries (SADC) to discuss training of those working in statistics offices, the training needs of South Africa were highlighted, against the background of an extreme shortage of official and other statisticians (14th Conference of Commonwealth Statisticians, Cape Town, September 2005). This related not only to professional statisticians, but users as well. It was decided that Pali Lehohla and myself would study the SADC documents on training, and present the South African (SA) position paper (Plenary session on 'Training of Statisticians in the SADC region,' 2005 conference of the SA Statistical Association, Grahamstown, November, 2005, attended by a large number of heads or representatives of statistical offices of the Southern African Development Countries and of university statistics departments, as well as general delegates). We indicated that, for SA, we needed training in statistics and statistical literacy among:

- Professional official statisticians, who plan surveys and analyze the results;
- Other professionals such as economists, demographers, geographers, social statisticians, and those setting up databases within the National Statistics System, who need to do further analyses;
- Policy makers and others who need to work with, and make sense of, data from Stats SA, in other government departments (national and provincial), municipalities, Non-Governmental Organization's, business, etc.; as well as to turn their "what I need to know," such as "are the poverty reduction programs working," into questions they can answer using the data;
- The general public as providers of statistical data – if they have some understanding of the basis for the allocation of funds to provinces and departments, and how this affects service delivery, we should get greater willingness to provide information. Such statistical literacy is also a pre-requisite for economic literacy, and as such is an essential skill for many South Africans, particularly those involved in wage negotiations, and those starting small businesses. It is also essential in preventing people being taken in by pyramid and other fly-by-night schemes, and the ubiquitous e-mail scams. In order to deal with HIV/AIDS

prevention and treatment issues effectively, they also need to be able to interpret statements concerning data made in the media on this subject.

- School children and teachers. In addition to making them statistically literate, we need to convince them that stats is fun. This should help ensure a supply of entrants to universities who want to work with statistics, especially official statistics. Also, if they have understanding, they can spread this to their families. If they enjoy collecting and looking at statistics, as a way to understanding the world they live in, we may be able to turn them to problem solving in maths (as a necessary evil if nothing else). This may assist in increasing the pool of mathematically literate students – SA has a big problem at present in that a number of students entering stats courses at universities with very good maths marks, have obtained those marks via a good memory, and practicing lots of past papers. They seem to work on the 'monkey see, monkey do' principle (ie if you see a phrase in a question, then this is the recipe to follow).

WHAT HAVE OTHERS SAID ABOUT THE PROBLEM AND POSSIBLE SOLUTIONS?

The issues have frequently been recognized and addressed: for example, Moore (1997) states, in the context of teaching major stream students, that:

> The trouble with probability is that it is conceptually the hardest subject in elementary mathematics. Psychologists, beginning with Tversky and Kahneman, have suggested that our intuition of chance profoundly contradicts the laws of probability that describe actual random behavior. They have also demonstrated that incorrect concepts remain firmly embedded in students who can correctly solve formal probability problems.
>
> We run the risk- no, we face the near certainty – that students will learn a formalism not accompanied by a substantial understanding of the behavior that the mathematics describes. Probability is the count of favorable outcomes divided by the count of all outcomes. Probability is area under a curve and can be found by integration. The record suggests that we are unlikely to move most students beyond that level of understanding.
>
> One root of the trouble with probability is lack of experience with the long-term regularity that the mathematics purports to describe. Chance variation is familiar, but chance appears haphazard because we very rarely see the large number of similar trials needed for the emergence of regular patterns. It is not accidental that games of chance, which impose a structure of repeated independent trials, were the historical setting for Pascal and Fermat and have been a staple of teaching ever since. Simulation allows learners to gain some experience with long-run chance behavior.
>
> Probability also illustrates the *power of abstraction* in mathematics.
>
> As George Cobb likes to say, 'In mathematics, context obscures structure. In data analysis, context provides meaning.'
>
> The trouble with statistics is that it is not mathematics.

For the above reasons, Moore (1997) suggests that probability be taught in a mathematics core, but recommends the use of simulation (e.g., coin tossing) to assist in teaching about long-run chance behaviour.

For service course students, theory is a minor consideration. These students are not interested in statistical theory (and are often antagonistic to and afraid of the subject and course). They need to understand the basics of statistical methods, for application in their major stream courses, in later years. In addition, Macnaughton (2002) notes that "The introductory course for non-statistics-majors is important because it is a main seedbed of public opinion about the field of statistics."

A multitude of other papers suggest a number of examples and aids that can be used to enhance teaching. To mention but a few: Borovcnik (2005) suggests that in order to illustrate probabilistic and statistical thinking, contextual examples are useful. Dixon (2002) looks at using "sample surveys as a vehicle for integration of statistical concepts" Redd (2003) argues (in the context to teaching stats to political science students) "that students must be taught how to think more critically about social and political issues." He notes that the lecturer of the course he was

subjected to was “one of the most interesting, insightful and humorous professors I’ve ever had,” which has led him to model his stats courses on the one he took. Berk (1996) evaluates the use of various types of humour, in examples, jokes, material in exams, etc.

HOW CAN ALL OF THIS BE ACHIEVED USING SOCIAL EXAMPLES?

Relating techniques to current social issues can assist in further breaking down the barriers, and increasing awareness of the usefulness of statistics. Using these examples can also enhance other aspects of what universities try to achieve, such as academic citizenship.

Service courses in stats are, at the University of the Witwatersrand at least, usually taught before students have any useful grasp of the areas in which they are going to apply these methods. This make it very difficult to use examples that interest them and that their lecturers think are sensible. The lecturers forget that the students are in 1st year, and have not yet grasped the details and implications of the basic concepts of the area… In addition, these service courses are often given to students from many branches of a disciple, such as “commerce” and “biological sciences.” Some progress can be made by pointing out that anyone who crosses a road, or drives a car, or follows various clubs/teams in sport is an expert in probability. (If they are in the lecture and crossed a road – they survived!) Examples of “student issues,” and weekly expenditure on soft drinks, can also be useful.

Using social examples, particularly concerning HIV/AIDS, in a country where this is a major problem, has many advantages. We can raise the issues of HIV/AIDS prevention and treatment from a number of different points of view:

- Using the issues/data as examples to illustrate theoretical aspects of stats: allows us to mention the subject without too much embarrassment among young students who are not convinced that lecturers know anything about the forbidden subject.
- Raising the issue of “where do we get the data from” (hence creating awareness of official statistics – both the collection of numbers, and the reasons why one needs to collect them, as well as different data sources), and “why do different data sources differ” (e.g., number of deaths from HIV/AIDS in the past years).
- Providing some background on the disease (e.g., the background of a CD4 count, in order to put data for a box and whisker or whatever plot into context).
- Discuss the ethical dilemma - issues of “what is best for me, what is best for the community,” by discussing various medical tests or treatments with different costs, and different efficacies – if I was a patient, which would I want; if I was a medical practitioner, which would I go for.
- By educating the students about HIV/AIDS in a statistical context, also helps to enlist our students in getting their fellow students to accept that there is a risk that needs to be dealt with. In addition, in SA, many of our students are the first from their home areas to come to a university, and as such have some standing in their community. As such we hope they will serve as a method of communication.

SOME SPECIFIC EXAMPLES CURRENTLY USED IN OUR COURSES

In an introductory talk to students wanting to do Actuarial Science and Mathematical Statistics at first year level, about what mathematical statistics is all about, I make the following points.

- Stats is about collecting and analyzing data, and using this to make inferences about what is happening now, and what is like to happen in the future.
- Ask questions as to what is happening in the world (e.g., AIDS, or whatever else is currently in the media), and expand on implications of these events.
- Why do we need to know this information? What are the implications as to planning?
- What are the issues involved – personal versus community? (e.g., one heart transplant versus anti-retrovirals for *x* people)

Actuarial Science 1:

- The demography section has a tutorial based on data from a Medical Research Council technical reports concerning "The Impact of HIV/AIDS on Adult Mortality in South Africa,”

as well as other questions relating to life table calculations for populations in the presence of HIV/AIDS.

Mathematical Statistics 1:

- Examples on the proportion of HIV positive individuals in a group to illustrate basic probability rules, conditional probability and Bayes' theorem.
- Examples in descriptive statistics.

Mathematical Statistics 2:

- Example of how to ask an embarrassing question in a survey ("ask" either the positive or negative form of the question randomly, so that the answer "True" may mean that the respondent has/has not the disease, without the questioner knowing which question was asked!). It was originally used to ask about illegal abortions, but students are invited to think of other embarrassing questions (or illegal activities) that this method may be used in order to encourage a (truthful) response. Drugs, homosexual activity, and of course HIV/AIDS are typical. Also examples such as possible transmission rates of HIV from sex workers, and transmission by mosquitoes.
- Examine the use of maximum likelihood in inferring the proportion of HIV positive individuals in a population and obtaining various confidence intervals for this proportion.

Actuarial Science 3:

- In the Survival Models topic, the problem of projecting mortality rates due to AIDS is considered. We are currently looking at spending more time on familiarising students with the AIDS models produced by the Actuarial Society of SA (ASSA).
- Computational component: Excel assignment based calculations and discussion related to articles concerning the rise in number of adult deaths in South Africa.

Mathematical Statistics 3:

- Use HIV/AIDS status: positive or negative, and then negative, positive - not symptomatic, positive-symptomatic, to teach indicator variable regression.
- Also use viral load, etc, in teaching variable transformations.
- Digress into other aspects of mother to child transmission, cost of testing versus cost of anti-retrovirals, etc.

Mathematical Statistics honours and masters by coursework and research report:

- Sampling course: sampling for estimation of HIV prevalence, differences in sources of data re deaths. Questionnaire design, contingency tables.
- Biostatistics course: Get people from Medical School to teach some of the medical background. For example, Mother-to-child transmission, general transmission, implications of different types of testing (and costs) as to whether baby is positive or negative (including concept of sero-reversion), individual versus populations issues in costs of tests and treatment.
- Official Statistics course: Discuss the implication on the country's demography, estimating HIV prevalence, ASSA model, implications on labour data, requirements for planning of hospitals, etc.
- Students are encouraged to take courses in Demography, where social issues are also addressed. (They are also encouraged to take a course in Geographic Information Systems, which provides a very powerful combination when combined with spatial statistics.)
- Honours project course – students are required to read and present papers (generally from *Significance*, else their own choice), which generally concern the social aspects of stats. The major project also generally concerns analysis of data, eg concerning HIV/AIDS studies done at Wits, predictors of university pass rates, etc.

Actuarial Science honours:

- Apart from the examples used in the courses (e.g., on Health Statistics, Insurance, Product Development), the students are also required to analyse data from Stats SA, combined with economic data from the Johannesburg Stock Exchange, and the All Media and Products Survey (AMPS) – a major market research study. This includes ways of identifying potential customers for insurance products, and characterising their demographics and behaviour.

Actuarial Science masters:

- Considers the impact that deaths from HIV/AIDS has on mortality of the lives that are effecting life assurance and related products. A number of papers and reports on the prevalence rates and progress with the syndrome are read and discussed. Knowledge of the regularly updated ASSA AIDS model is required. They are also required to consider what mortality tables they should use in designing products and valuing actuarial liabilities of a South African life company.

Service courses to commerce and biological sciences students:

- We have introduced CAI, and require completion of certain quizzes as a DP requirement. We are trying to bring in HIV/AIDS examples here.

CONCLUSION

The training of teachers in statistics, started at the ICOTS-6 conference, has continued, with a number of further courses being given at conferences etc. The teachers trained at ICOTS-6 have also given many courses to their colleagues. We are currently putting in place a more structured program to reach teachers in all areas. As the school syllabus now requires students to look at social issues, such as crime trends in their neighbourhood, we will be spending time looking at where to source such data. We are planning another *CENSUSATSCHOOL* to provide further data for the classroom, and distributing Stats SA paper and electronic products for use.

Training at universities – after the SASA conference, we are rolling out a program aimed at including material related to official statistics in all university statistics departments. Several of the universities (and in particular the University of Cape Town) already use the electronic data from Stats SA for teaching – not just how to draw a histogram or a obtain a random sample, but deciding on what this means.

Assessment of which techniques/ examples work in the Wits courses? This is still work in progress. We have not had time to quantify the usefulness of these examples as yet (December 2005). Hopefully by the ICOTS conference in June …

REFERENCES

Berk, R. A. (1996). Student ratings of 10 strategies for using humor in college teaching. *Journal on Excellence in College Teaching*, 7(3), 71-92.

Borovcnik, M. (2005). Probabilistic and Statistical Thinking, http://cerme4.crm.es/Papers%20definitius/5/Borovcnik.pdf.

Dixon, P. (2002). Sample surveys as a vehicle for integration of statistical concepts. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.

Macnaughton, D. B. (2002). The introductory statistics course: The entity-property-relationship approach, http://www.matstat.com/teach/eprt99.pdf.

Moore, D. S. (1997). Probability and statistics in the core curriculum. In J. Dossey (Ed.), *Confronting the Core Curriculum*, (pp. 93-98). Washington, DC: Mathematical Association of America.

Redd, S. B. (2003). Rejecting the dull hypothesis: Making undergraduate research methods courses interesting. Paper presented at the annual Meeting of the Midwest Political Science Association, April 2-6, Chicago, Illinois.