Using Real Life Examples to Teach Abstract Statistical Concepts

KEYWORDS:
Teaching;
Conceptual understanding;
Statistical competency;
Statistical literacy.

Nyaradzo Mvududu
Seattle Pacific University, Washington, USA.
e-mail: nyaradzo@spu.edu

Gibbs Y. Kanyongo
Duquesne University, Pittsburgh, Pennsylvania, USA.
e-mail: kanyongog@duq.edu

Summary
This article provides real life examples that can be used to explain statistical concepts. It does not attempt to be exhaustive, but rather, provide a few examples for selected concepts based on what students should know after taking a statistics course.

◆ INTRODUCTION ◆

Efforts in statistics education reform have focused on statistical thinking and conceptual understanding rather than mere knowledge of procedure. Statistics is not an easy subject to teach. Almost every statistics beginner experiences difficulties understanding the topic. There is a consensus among statisticians that statistics education should focus on data and on statistical reasoning rather than on either the presentation of as many methods as possible or the mathematical theory of inference.

Generally, the goal of statistics education is to answer 'real world' questions. The student should develop sufficient competence to understand and draw accurate meaning from a statistical argument. Examples therefore, should be presented in the context of real world problems. Recommendations for statistics courses from the GAISE (Guidelines for Assessment and Instruction in Statistics Education) Report (Garfield et al. 2005) include the following:

- Emphasize statistical literacy and develop statistical thinking.
- Use real data.
- Stress conceptual understanding rather than mere knowledge of procedures.
- Foster active learning in the classroom.
- Use technology for developing conceptual understanding and analysing data.
- Use assessments to improve and evaluate student learning.

Furthermore, the guidelines given in the GAISE Report suggest designing statistics courses that teach students to communicate results of a statistical analysis and do so in context. In addition, students should develop skills to read and critique news stories and journal articles that include statistical information. That is, they should have statistical reasoning skills as well as statistical understanding. This goal of applying statistics to everyday life is reiterated by Gal and Ginsburg (1994).

It is sometimes difficult for students to relate statistical concepts presented in class to real world problems and everyday situations. Therefore, teaching introductory statistics requires the instructor not only to transmit knowledge, but also to enhance students' motivation and attention (Symanzik and Vukasinovic 2006). From our experiences in teaching statistics, what we have found helpful is to relate the statistical concepts to real life situations. In most cases, this strategy helps lower the students' statistical anxiety. The impact of statistics anxiety on student learning is well documented (e.g. Benson 1989; Dillon 1982; Roberts and Bilderback 1980; Roberts and Saxe 1982).

Researchers (e.g. Everson et al. 2008) have provided some ideas for putting the GAISE guidelines into practice in the statistics classroom. Interactivity, hands-on exercises, visualization of statistical concepts and well-documented real life examples are some of the features of a statistical course that help stimulate the student's activity in class, ease understanding of statistical concepts and make the whole course more attractive and fun for the student.

In this article, we discuss certain real life examples that we normally use to explain statistical concepts. We do not attempt to cover all the statistical concepts, but rather, we just provide a few examples that are popular with students. Each of the concepts discussed can be linked directly to the goals suggested in the GAISE Report.

◆ REAL LIFE EXAMPLES ◆

If the field of statistics could be described in one word, that word would be *variability*. In almost all cases we are attempting to explain variance or apportion variance. We use models (samples) as representation of the real world so it should not be surprising that there is variability in whatever characteristics we study. As the GAISE Report suggests, students should understand that variability is natural and is also predictable and quantifiable.

The weather reports we hear on the news everyday provide an illustration of the value of variability. While two regions can have the same average temperature of, say, 55°F, we get a better understanding of the weather in these regions by considering the variation in temperature. For instance a region that is near the coast is more likely to have a narrower range than one further inland. That is, the average temperature of 55°F for the inland region may have highs of 75°F and lows of 35°F, while in the coastal region the highs might be 60°F with lows of 50°F. Although the regions have the same average temperature, the higher standard deviation means that predictions are less reliable for the temperature of the inland region.

A similar phenomenon can be found in sports. In sports, trying to predict which teams will win, on any given day, may include looking at the standard deviations of the team ratings in various categories. In such ratings, anomalies can match strengths versus weaknesses in an attempt to understand which factors are the stronger indicators of eventual scoring outcomes. For instance, there will be teams that excel in some aspects of the sport and perform poorly in others. Teams may have a great offence and a weak defence or vice versa. Furthermore, teams with higher standard deviations will be more unpredictable. It is important, of course, to consider the mean with the standard deviation in making predictions. A team that is constantly good in most categories will have a low standard deviation but so will a team that is consistently bad.

Candidate	*Polling agency*	
	CNN	*Zogby*
John McCain	44%	42%
Barack Obama	44%	45%
Others	7%	6%

Table 1. Polling results for the 2008 US presidential race, 23–24 October 2008

It is important that students understand the concept of a *sampling distribution* and how it applies to making statistical inferences. Sampling distributions are important because inferential statistics is based on them. Inferential statistics is about drawing conclusions about the population based on sample data.

An example of this is the polling done prior to elections. At the end of August 2008, two of the polls conducted to predict the winner of the US presidential race provided somewhat different results (see table 1). Looking at these results provides an opportunity to discuss what happens when different samples are drawn from the same population. Although the samples may have been drawn from the same population they may provide different estimates. Furthermore, individual sample statistics do not always match the true population value, but vary around it. At this point the instructor can discuss the distribution of sample statistics (e.g. means being approximately normal).

This demonstration can also naturally lead to a discussion of the margin of error as it is computed in political polls and to confidence intervals (Baumann and Danielson 1997). Examining the formula for the margin of error can also lead to a discussion of the effect of sample size and variability in the sampling statistics. Students can see that larger samples correspond to smaller sampling variation and smaller margins of error, and that population size is not a factor in the margin of error. With these ideas clarified, students can turn to applications such as projecting winners in elections or comparing the responses of subgroups such as men and women.

In *hypothesis testing* students learn how to make appropriate use of statistical inference. It is impor-

Hypothesis test	Jury trial
Null hypothesis: The researcher assumes that the treatment is not effective unless there is compelling evidence showing otherwise.	In the US criminal justice system, it is assumed that the defendant is innocent until the prosecutor presents compelling evidence showing otherwise.
Test of significance: The researcher is confident that the treatment is effective when it is very unlikely that the data could occur simply by chance.	The jury returns a 'guilty' verdict on the defendant when they are convinced beyond any reasonable doubt by the evidence.
Data: The researcher presents the data collected to help demonstrate that the treatment is effective.	The prosecutor presents incriminating evidence to link the defendant to the crime.
Critical region: A test statistic computed from the data either falls in the critical region – in which case the researcher has enough evidence to reject the null hypothesis, or does not fall in the critical region – in which case there is not enough evidence to reject the null hypothesis.	The evidence presented is either compelling enough to convince the jury that the defendant is guilty, or the evidence is not that compelling.
Conclusion: If the test statistic is not in the critical region, the researcher 'fails to reject the null hypothesis'. This does not mean the null hypothesis is true, it just means there was not enough evidence to reject it.	If the evidence is not compelling enough to convince the jury, the jury recommends 'acquittal'. This does not mean the defendant is innocent, it just means the evidence was not compelling enough to recommend 'conviction'.

Table 2. Hypothesis testing and jury trials compared

tant that students realize that whatever the conclusion, they cannot prove anything, but rather provide evidence to support a certain conclusion. It is up to the researcher to show that an effect is unlikely to have occurred merely by chance. This is similar to the legal system where the burden of proof is on the prosecutor (see table 2). Television dramas such as *CSI: Crime Scene Investigation* (http://www.cbs.com/primetime/csi/) illustrate this point. The defendant is presumed innocent until the evidence suggests otherwise. This is a good point to introduce the idea of the null hypothesis. At what point is there sufficient evidence to conclude that the defendant is guilty? This question implies that sometimes the evidence is not enough to convict someone. From here the instructor can discuss alpha levels and *p*-values. Table 2 helps demonstrate the parallels between hypothesis testing and the criminal justice system.

Statistics requires data collection and recording. It is important to consider how data are collected and recorded to come to valid and reliable conclusions. *Measurement*, if not done properly, can introduce bias in surveys and experiments.

The New York Times (Schwarz 2007) carried a story alleging racial bias in how the National Basketball Association (NBA) referees call fouls on players. The fouls were measured in terms of fouls-per-minute rates of the players. Because data on calls per specific official were not available to the researchers, the number of calls was measured per crew of three officials. The NBA commissioner disputed the study as flawed partly because of the way variables were measured. He stated that, 'I don't believe it. I think officials get the vast majority of calls right. They don't get them all right. The vast majority of our players are black.'

This story provides an opportunity for a statistics instructor to discuss the impact of measurement on statistical analyses conducted and conclusions drawn from a study. Additionally, the class can discuss factors that can have an effect on measurement and some ways to address potential bias from such factors. This can lead into a discussion of validity and precision in research studies.

In statistics and research we talk of 'significant' findings. It is important for students to understand what we mean by this. *Statistical significance* does not mean the same thing as profundity in findings. The word 'significant' takes on a different meaning from the common usage of the word. Furthermore, statistical significance does not necessarily imply practical importance. In the case of alleged racial bias in fouls called by the NBA referees we might consider the size of the difference. If the difference is, say, about one extra foul per game we might ask the question, 'Does it affect the game?'

The class can discuss how it is that we can have statistical significance and very low practical significance. This is an opportunity to discuss factors that can impact significance such as sample size. It can lead to a discussion of *type I* and *type II errors*. Students can think about the implications of committing one type of error over the other.

When we conduct a hypothesis test we prove neither the null hypothesis nor the alternative hypothesis. Rather we provide evidence to support one conclusion or another. Whichever conclusion we reach, there is a chance that we are wrong.

Using the legal system example, maybe we just happened to find incriminating evidence when the defendant is really not guilty. In this case, we have committed a Type I error. Think of it as a false alarm or a false positive. On the other hand, let's say the defendant was a 'good' criminal and did not leave much evidence behind. We conclude that he is not guilty when in fact he is. We have committed a type II error. We have failed to see something that is there. It is important to consider the consequences of committing each type of error.

Correlation analysis, a procedure in testing hypotheses, is widely used in statistics as a means to establish if variables are significantly related. A common error in interpreting the results of a correlation analysis is to assume that the existence of a statistically significant association is evidence of a causal relationship. The following example illustrates the fallacy of such thinking. In an effort to reduce greenhouse gases, the Japanese environment minister instituted a policy that required businessmen to stop wearing suits and ties (Kestenbaum 2007). It was reported that this action led to a reduction of two million tonnes of greenhouse gases from the country's growing emissions. The title of the radio programme was *Japan Trades In Suits, Cuts Carbon Emission*. This provides an opportunity for the instructor to discuss the fact that the two variables may both be related to a third variable. In this case the third variable was the level of air conditioning. While two variables may have a strong correlation, we cannot draw the conclusion that one causes the other.

◆ CONCLUSION ◆

In addition to using real world examples, some of the ideas suggested to facilitate the learning of statistics include making the content relevant to the students (e.g. Chance 2002), using humour (e.g. Friedman et al. 2002), and using projects (e.g. Fillebrown 1994; Smith 1998). All these approaches produce beneficial outcomes. In this article we chose to focus on the aspect of real life examples for a small selection of concepts.

The ultimate goals in our classes: to develop statistical literacy and competency in our students. Quite often students will ask, 'Why am I taking this course?' Participation in the course should lead them to answer that question with, 'Because data are interesting and useful in understanding the world.' As statistics deals with uncertainty in the real world we teach our students caution in drawing conclusions from statistical analyses. In particular, we think it is important our students approach questions from multiple perspectives. By teaching our students in this fashion we believe we are providing them the tools necessary to develop statistical literacy and competency.

References

Baumann, P.R. and Danielson, J.L. (1997). Statistics in political polling. *Stats: The Magazine for Students of Statistics*, **20**, 8–12.

Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education*, **57**(3) 247–261.

Chance, B. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, **10**(3), http://www.amstat.org/publications/jse/v10n3/chance.html (accessed January 18, 2007).

Dillon, K.M. (1982). Statisticophobia. *Teaching of Psychology*, **9**(2). 117.

Everson, M., Zieffler, A. and Garfield, J. (2008). Implementing new reform guidelines in teaching introductory statistics courses. *Teaching Statistics*, **30**(3), 66–70.

Fillebrown, S. (1994). Using projects in an elementary statistics course for non-science majors. *Journal of Statistics Education*, **2**(2), http://www.amstat.org/publications/jse/v2n2/fillebrown.html (accessed November 21, 2005).

Friedman, H.H., Friedman, L.W. and Amoo, T. (2002). Using humor in the introductory statistics course. *Journal of Statistics Education*, **10**(3), http://www.amstat.org/publications/jse/v10n3/friedman.html (accessed January 16, 2006).

Gal, I. and Ginsburg L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, **2**(2), http://www.amstat.org/publications/jse/v2n2/gal.html (accessed November 21, 2005).

Garfield, J., Aliaga, M., Cobb, G., Cuff, C., Gould, R., Lock, R., Moore, T., Rossman, A., Stephenson, R., Utts, J., Velleman, P. and Witmer, J. (2005). *Guidelines for Assessment and Instruction in Statistics Education (GAISE)*. http://www.amstat.org/education/gaise/ (accessed January 16, 2006).

Kestenbaum, D. (2007). Japan trades in suits, cuts carbon emission. *Morning Edition, National Public Radio*. http://www.npr.org/templates/story/story.php?storyId=14024250 (accessed October 1, 2007).

Roberts, D.M. and Bilderback, E.W. (1980). Reliability and validity of a statistics attitude survey. *Education and Psychological Measurement*, **40**(1), 235–238.

Roberts, D.M. and Saxe J.E. (1982). Validity of a statistics attitude survey: A follow-up study. *Educational and Psychological Measurement*, **42**(3), 907–912.

Schwarz, A. (2007). Study of NBA sees racial bias in calling fouls. *The New York Times*. http://www.nytimes.com/2007/05/02/sports/basketball/02refs.html (accessed February 5, 2007).

Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education*, **3**(3), http://www.amstat.org/publications/jse/v6n3/smith.html (accessed November 21, 2005).

Symanzik, J. and Vukasinovic, N (2006). Teaching an introductory statistics course with CyberStats, an electronic textbook. *Journal of Statistics Education*, **14**(1), http://www.amstat.org/publications/jse/v14n1/symanzik.html (accessed October 1, 2007).

STATISTICAL QUOTES

"By a small sample we may judge of the whole piece." – Miguel de Cervantes, *Don Quixote*, 1605

"Data! Data! Data!" he cried impatiently. "I can't make bricks without clay." – Sherlock Holmes to Dr Watson in *The Adventure of the Copper Beeches* by Sir Arthur Conan Doyle, 1892

"Statistics means never having to say you're certain." – Myles Hollander

"It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge." – Pierre-Simon Laplace, *Théorie Analytique des Probabilités*, 1812

"It is no great wonder if in long process of time, while fortune takes her course hither and thither, numerous coincidences should spontaneously occur." – Plutarch, *Plutarch's Lives: Vol. II, Sertorius*, 75 AD

"Coincidences, in general, are great stumbling blocks in the way of that class of thinkers who have been educated to know nothing of the theory of probabilities – that theory to which the most glorious objects of human research are indebted for the most glorious of illustrations." – C. Auguste Dupin in *The Murders in the Rue Morgue* by Edgar Allan Poe, 1841

"That the chance of gain is naturally over-valued, we may learn from the universal success of lotteries." – Adam Smith, *The Wealth of Nations*, 1776

"The great body of physical science, a great deal of the essential fact of financial science, and endless social and political problems are only accessible and only thinkable to those who have had a sound training in mathematical analysis, and the time may not be very remote when it will be understood that for complete initiation as an efficient citizen of one of the new great complex worldwide States that are now developing, it is as necessary to be able to compute, to think in averages and maxima and minima, as it is now to be able to read and write." – H.G. Wells, *Mankind in the Making*, 1903