PREFACE TO THE PROCEEDINGS OF THE 2012 IASE ROUNDTABLE ON TECHNOLOGY AND STATISTICS EDUCATION

Robert Gould
Dept. of Statistics
University of California, Los Angeles, USA
rgould@stat.ucla.edu

This 2012 Roundtable was convened to discuss issues surrounding the use of Technology in teaching statistics. This was not the first Roundtable on this topic. Indeed, it's interesting to compare the papers in these *Proceedings* with those in from the 1996 Roundtable held in Granada, Spain. (http://www.stat.auckland.ac.nz/~iase/publications.php) Technology has changed the educational landscape dramatically since 1996, but many of the same topics appear in both Proceedings: the use of graphing calculators, the role of statistical software, the possibilities of distance learning. The 2012 Proceedings highlight some of the particular issues and challenges facing us with the current state-of-the-art.

The role that mathematics should play in statistics education played perhaps a more prominent role in the 2012 Roundtable than in 1996, I think because the power of technology to replace mathematics as a tool for understanding is becoming more apparent. Historically, at most institutions around the world, I think it's safe to say that the responsibility for teaching statistics has largely been assigned to mathematicians. Indeed, even when taught by statisticians, there is still a struggle over the balance between mathematical and statistical content. In his 1940 address to the Institute of Mathematical Statistics, Harold Hotelling outlined a course of statistical education that was quite influential in the U.S. for the next 50 years (Hotelling, 1940). (And perhaps still very influential in places.) While Hotelling believed that some college students could get by without a calculus-based statistics class, he believed that students who needed to work with data did, in fact, need to first master calculus and perhaps even more advanced mathematics.

The word calculus is derived from the Latin word used to describe stones or pebbles used for counting and calculating. (Oxford English Dictionary, www.oed.com.) So the calculus is a method of computation and, more importantly, a method that provides tools to help us think. For over a century calculus has been an important tool for how we think about statistics.

As George Cobb put it in his paper "The Introductory Statistics Course: A Ptolemaic Curriculum", the logic of statistical inference has been shaped by the "Tyranny of the Computable." (Cobb, 2007). In other words, we have been limited by what we could compute. Much of mathematical statistics consists of approximations and large-sample limits that help us find approximation solutions when exact solutions are too difficult to compute. Much of statistics education has been spent helping students learn how to apply these approximations. When teaching introductory statistics, the challenge has been to motivate these approximate methods without direct reference to the underlying mathematics, which is a particularly difficult challenge if one's goal is teach statistical thinking and reasoning.

But modern computers have made very fast calculations possible and now allow us to find exact calculations in a way that would have made Fisher weep for joy. Indeed, calculations are so fast that modern statistics is now inseparable from the computer. And this presents many new challenges to educators as some topics become a more important part of the curriculum while other topics become obsolete. This Roundtable is one venue for discussing this challenge.

Computers do much more than simply speed up computations; they are now one of the primary mediums for entertainment. And tools used to entertain can also be used to educate. The ability to create and edit videos, to foster interactivity between students and the material, to integrate multi-media with traditional classroom lectures-- these abilities have provided educators with an embarrassment of riches for teaching statistics. We've had centuries of experience with

books and chalkboards, but we are only beginning to understand how students and teachers interact with computers. We are now challenged to incorporate discoveries concerning psychology and cognition so that we can understand how to use these tools. And, so often what we can do with technology is beyond what we know about how the mind works with and perceives the technological tools we create.

A wonderful surprise brought by the computer (a surprise to me, at least) is a new era: the Age of Data. Data sets are growing larger and larger (some are growing larger as you read this) and now large and rich data are everywhere, many available to anyone with an internet connection. Once, students learned about data perhaps for the very first time when they took a statistics or science course. Now, one can argue that data make the world go round. Our global culture is built on an infrastructure that moves data around the world, and collects data about everything under the sun. The challenge here is to understand these oceans of data. And the challenge facing this Roundtable is how to prepare our students to face these data.

The change in the scale and availability of data has brought about a change in the way our discipline interacts with others. In the past, a researcher brought data to a statistician with a specific question in mind. And the statistician analyzed the data so as to provide an answer. This was because data were collected deliberately and with purpose. (In more extreme cases, statisticians collaborated with researchers to assist in the data gathering and to be on hand to resolve any hypotheses which might be posed.) But today, data are often collected by machines, without direct human intervention. These machines are programmed with some vague intent, perhaps so that a grocery store chain can keep track of inventory, or simply because a government agency requires that information be recorded; these data often provide so much more information than the designers originally needed that the only question asked of the statistician is "what can we learn from these data?" This places the statistician very much in the role of a scientist. To me, this is one of the great challenges in teaching future data scientists: how do we teach students to engage with data with the curiosity of a scientist?

A very important change brought on by fast computing is universal statistics education. In 1940, Hotelling mentioned that the development of statistical theory had led to an amazing increase in the demand for statistics and for statisticians. And with this increase came a demand to educate and teach more statisticians. Hotelling believed that all students should learn statistics, but for him 'all students' meant all college students. Today we are again facing an amazing increase in the demand for statistics and statisticians. This demand was created not by theory, but by technology. And today we recognize that statistics must be taught not only to college students, but to students of all ages, regardless of their future careers. However, very few of our children's teachers have themselves studied statistics, and so many of us our challenged to find ways to teach teachers to teach statistics.

Added to these challenges is the very important fact that technology is expensive. Not only is there great expense on individual students and teachers, but great expense on schools and on governments to maintain and build the infrastructure that modern statistics relies upon. These challenges are felt particularly strongly in developing countries, and also to varying degrees in some sections of the richest countries. One of the challenges we will consider this week is very basic. If we believe that all students must learn statistics, how can we make sure that technology helps makes this possible, and does not stand in the way of quality education for all?

Modern computers have brains of silicon, an element that is commonly found in some form in dust and sand, and this observation brings us back to calculus. We are still calculating with small pebbles, only now the pebbles are very, very small. The papers in this volume of Proceedings confront the challenges that are produced by this new calculus. The computer requires us to teach more students at more levels, in all countries around the globe. We are challenged to use the computer to develop deeper student understanding because modern life requires a deeper understanding than at any other time in history. We are challenged to develop new tools and methods so that students and others can interrogate rich and complex data.

This Roundtable gathered researchers and statistics educators from ten different countries for a week in Cebu City, The Philippines, to present and discuss their responses to these challenges. The discussions were engaging and productive and interesting. These Proceedings contain the papers as they were presented (although some minor changes were made.) All papers were peer reviewed in a blinded review before acceptance to the Roundtable. Later this year, a peer-reviewed special issue of Technology Innovations in Statistics Education (tise.stat.ucla.edu) will provide these researchers with the opportunity to integrate the discussions and comments made during the Roundtable into their original papers.

Grouping these papers into categories was difficult, since all of these papers defy such simple summarizing. Still, as an editor I felt it an obligation to force papers into some sort of slot, even if the authors might find the fit uncomfortable. And so, we have three papers that describe exciting new tools for teaching statistics. The Island (Bedford, Baglin, and Bulmer) provides a virtual environment that greatly extends the possibilities for teaching advanced statistical topics in a constructive, student-centered environment. Data Games (Erickson) provide new possibilities for motivating students by showing them that statistics can improve their lives (or at least their scores on games). Budgett, Pfannkuch, Regan and Wild offer new and innovative visualization tools and demonstrate their effectiveness at developing conceptual understanding.

Three papers address the challenges of giving students access to large, complex, and real data. Finzer discusses the need for re-thinking statistics education in a multi-disciplinary framework (a theme that many of the authors discuss) if we are to prepare future data scientists. In a pair of papers, Ridgway, Nicholson, and McCusker discuss new tools and new frameworks to help both teachers and students to access, analyze and understand multivariate relationships in the context of the 'semantic web', which is particularly timely since more data seem to be housed 'in the cloud'.

Using technology to improve statistics education in developing countries was a theme of particular prominence this year. Two papers addressed educational innovations in Kenya (Stern, and Stern, Coe, Stern and McDermott.) And Krishna & Idris remind us the graphing calculator is a prominent educational tool throughout the world, and discuss ways it is used in Malaysia.

Finally, a trio of papers discuss the use of technology in the classroom. Reston points out that we often implement technology without a well-defined theoretical framework, and provides us with a framework that has been helpful in redesigning statistics instruction in the Philippines. Hassad reports on a study of the attitudes of college-level instructors on using technology in their statistics classrooms. And Mocko describes the wide variety of tools available for teaching on-line statistics courses, as well as explaining which tools seem to work best in at least one course offered in the United States.

REFERENCES

Cobb, G.W. (2007). The introductory statistics course: A ptloemaic curriculum? *Techol. Innovat. Stat.Educ*., **1**(1), Article 1, Available at http://repositories.cdlib.org/uclastat/cts/tise/vol1/iss1/art1.

Hotelling, H. (1940). The teaching of statistics. *Ann. Math. Statist.*, **11**(4),457-470.