

Examining Data Collection and Criteria Selection in Introductory Statistics

Jeffrey Clark
Elon University
Joint Mathematics Meetings
January 18, 2003

Elon University is a liberal arts school with roughly 4000 students and is located in central North Carolina. *General Statistics* is a required course for incoming freshman.

Introduction

I have taught Elementary Statistics at Elon University for the past thirteen years. The students that take this course in the past have mostly come from the social sciences. This academic year Elon has begun to require all of its students to take either Elementary Statistics or Calculus I, so we are seeing roughly 80% of the freshmen in our classes.

Problem

We would like our students to be good consumers of statistics; for that fraction of our students who will become producers of statistics we also wish to give them a firm foundation in basic statistical techniques.

Unfortunately, being a good consumer means understanding the problems inherent in not only data analysis but also in data collection and even in the formation of statistical questions. A valid statistical analysis performed on biased data is useless. Even worse, a valid statistical analysis performed on a vague or misleading hypothesis is dangerous.

Existing Texts

The following texts are good introductions to statistical technique, but do not spend much time on the validity of data or the formation of statistical questions.

Workshop Statistics by Rossman et al spends two chapters out of twenty-seven on sampling technique. It does try to stress alternative explanations in its problems.

Triola's *Elementary Statistics* spends a couple of sections in the first chapter discussing these issues.

The History of Intelligence Testing

The late Stephen Jay Gould's book *The Mismeasure of Man* examines the history of intelligence testing, starting with Binet and leading up to the past history. The focus of the book is on the mis-use and abuse of statistics in the search to prove that white males are the smartest creatures on Earth. It is not an attack on intelligence testing so much as an attack on bad statistical practices, almost none of which could be found in a standard statistics textbook.

Fundamentally, intelligence is a fuzzy concept with many different components. Testers need to be very careful about what they claim a test measures.

In addition, if you start with a foregone conclusion (that white males are smarter), then your choice of criteria for intelligence is affected. Gould cheerfully describes the progression from (a) associating intelligence with larger cranial volumes until skulls of people of African descent were found to be larger on average; (b) associating intelligence with smaller cranial volumes until skulls of people of Asian descent were found to be smaller on average; finally to (c) associating intelligence with moderate cranial volumes.

Also, whenever measurement standards are the least subjective, it is vital to use blind experiments where the person taking the measurement is isolated from the knowledge of what group the subject is from. The use of compressible material such as seed to measure cranial volume fluctuated greatly, depending on whether or not the person measuring the cranial volume knew if the skull was Caucasian, etc.

If the tests produce a huge mode, as when tests were poorly explained to subjects who were not native English-speakers, then the data is questionable.

Alternatives

Huff's *How to Lie With Statistics* is a classic work on statistical skepticism.

Moore and McCabe's *Introduction to the Practice of Statistics* spends a good deal of time discussing good data and good hypotheses.

Conclusion

If our primary goals in our general statistics courses is to produce informed consumers of statistics, then we should supplement the statistical theory in most of our texts to also focus on bad statistics. Bad statistics includes not just biased sampling techniques but also the choice of criteria to support a foregone conclusion and the poor formation of some statistical problems from the start.

This presentation is available at:

<http://frodo.elon.edu>

under the Presentations link.