

Computer-Aided Statistical Instruction—Multi-Mediocre Techno-Trash?

C. J. Schwarz

Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6. E-mail: cschwarz@stat.sfu.ca

Summary

The very soul of statistics are data, but few students actually collect data as part of their statistical journey. The impediments to real data collection exercises are very real—they are logistically difficult to set up, expensive, and may not work because of extraneous events outside the control of the instructor. Computer-aided laboratories are a way to bring many of the benefits of actual data collection to students at a fraction of the cost and can be easily controlled by the instructor.

There are many computer-aided modules available—indeed a search on Google gave over 1 million hits. Some modules are good but many are mediocre. What separates the gems from the trash?

Key words: Computer simulation; virtual experiments; virtual surveys.

1 Introduction

Students learn statistics by doing statistics (Smith, 1998; Rossman & Chance, 1999). For students in introductory courses, this had led to much emphasis on hands-on activities where students collect data through simple experiments (e.g. Jowett & Davies 1960; Mackisack, 1994; Martinez-Dawson, 2003; Schaeffer *et al.* 1996). However, hands-on activities are difficult to implement for large classes, can be expensive in terms of supplies, and may not work well because of uncontrolled variations in the actual activities. Computer-aided instructional modules potentially offer a low-cost controllable medium to help students “do statistics”.

Many statistical educators (including myself to some extent) are stuck in the old paradigm where all students are assigned the the same exercises with the same pre-existing, pre-run experiments and the same set of data. This old paradigm, while suitable for some disciplines is particularly unsuited for teaching statistics because:

- Students may believe that there is one, unique answer to a scientific problem.
- Students do not see the extent that natural variation can change results when the identical experiment is repeated. For example, while they may compute a standard error for an estimate, they do not see that estimates do indeed vary among their fellow students.
- Students may believe that results are clear-cut because everyone always gets the same answers.