SOME EXPERIENCE COMPARING SIMULATION AND CONVENTIONAL METHODS IN AN ELEMENTARY STATISTICS COURSE

Paul Alper and Robert L. Raymond, University of St. Thomas, USA

This paper deals with several years of experience teaching an elementary course where the conventional, formulaic results are continually compared with the simulation results. The conventional is done primarily through Minitab rather than substitution into formulas; the simulation starts early in the course and proceeds throughout the term. The computer software, Resampling Stats, was one of the vehicles used as was Mintab's primitive simulation capabilities.

BACKGROUND

It is hardly a secret that elementary statistics is one of the least liked experiences for undergraduates. The conventional and all-too-typical approach is to resort to arcane formulas which are simultaneously boring and terrifying to the student. The ubiquity of the computer has made it possible to alleviate some of the pain but at the expense of adding yet another layer of mystery. In order to make the subject of statistics more interesting and meaningful, we have focused on simulation. Simulation may help the student understand the concepts, but unless the usual topics and formulas are discussed, the student will lack the vocabulary to converse with others now, in subsequent courses, or on the job.

We have now had several years of experience teaching an elementary course where the conventional, formulaic results are continually compared with the simulation results. The conventional is done primarily through Minitab rather than substitution into formulas. The simulation starts early in the course and proceeds throughout the term. The first simulations are Monte Carlo and deal with populations for which everything is known such as the binomial and Bernoulli; next, the central limit theorem is illustrated.

We have extended the simulation to the inferential part of the term, which is, of course, the most difficult for students. Here, bootstrapping and approximate randomization are used to obtain confidence intervals and significance tests for the mean of a population, the mean of the difference, the difference of means, difference in proportions, regression coefficients, regression response and chi-square testing; in short, roughly the entire set of problems typically considered conventionally, all the while comparing the results to the conventional which Minitab has printed out. The computer software, Resampling Stats (Simon, 992) was one of the vehicles used as was Mintab's simulation capabilities.

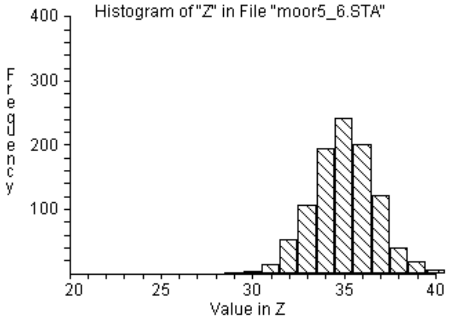
SOME HISTORY

Before embarking on the details of what we are doing and how it is evolving, it is well to reflect on a bit of history. Between the two of us, we have had almost a half century of experience teaching elementary statistics. Through the years, we have witnessed changes in the way the material is taught and presented. When we began, computers were few and far between with the result that emphasis was often placed--some would contend, MISplaced--upon successfully choosing a magic formula and plugging into it; such time-consuming activity left no time to discuss what the answer implied. Along came command-driven software which promised to relieve the agony of mindless and possibly incorrect calculations. Personal computers made possible menu-driven software and high-resolution graphics thus affording time to concentrate on what the results purportedly are telling us.

Unfortunately, the promise of conventional statistics software is largely unfulfilled. Students may be more comfortable pulling down a menu than using a formula; they may be more likely to pull down the correct menu than to substitute numbers into a formula successfully, but it is still overwhelmingly evident that for the majority of the students we face in an elementary statistics class, the subject is still viewed as fear-inspiring and largely incomprehensible. Succinctly put, the inferential part of the typical course--the hardest part of the course--has in it a cascade of logic such that the formulas are virtually unfathomable. For example, the t statistic and its methodology must be taken entirely on faith and it is no wonder that students fail to resonate either to a formula as it appears in a textbook or to a *deus ex machina* of a pull-down menu.

While an increasing number of instructors are utilizing simulation, they often stop at the central limit theorem and fail to use simulation in the inferential part of the course[a similar phenomenon occurred with the introduction of EDA: the first textbooks would introduce a few EDA techniques such as box plot and stem plot early on and then never mention them again when it came to inference]. We have taken the tack that simulation should be emphasized throughout the course especially now that bootstrapping--which when first brought forward had the appearance of black magic--has been legitimized.

Space does not permit illustrating a simulation for each and every typical problem in an elementary statistics course so here are merely two very typical ones, a univariate confidence interval problem and a so-called 2-sample t problem(*but without the need for the annoying t*). It should be noted that virtually any problem done in the course has a similar structure: the data is entered, a loop of 1000 repetitions is performed in order to obtain the resamples of the original sample and then some summary description is done after the loop. Both examples are taken from Moore's book, *The Basic Practice of Statistics* (Moore, 1995).

<table>
<tr>
<td>This pertains to problem 5.6 in Moore,
page 337 dealing with test scores
Read file "moor5_6.dat" a
repeat 1000
  sample 44 a b
  mean b c
  score c z
end
histogram z
mean z avg_z
stdev z stan_z
print avg_z stan_z
percentile z (2.5 97.5) mid95
print mid95
AVG_Z = 35.108
STAN_Z = 1.6615
MID95 = 31.977 38.273


</td>
<td>This is from Moore, problem 6.35,
page 449 and is solved via bootstrapping
This 2-sample example is from Moore,
problem 6.35, page 449. It deals with
the weight gain of chickens who
ate the experimental corn vs.
the weight gain of chickens
who ate the control corn.
Bootstrapping is used to solve
this Fisher-Behrens' problem.
Read moor6_5.dat" cntrl exprmtl
repeat 1000
  sample 20 cntrl a
  mean a avg_c
  sample 20 exprmtl aa
  mean aa avg_e
  subtract avg_e avg_c k
  score k z
end
histogram z
mean z avg_z
stdev z stan_z
print avg_z stan_z</td>
</tr>
</table>

```
MTB > TINTerval C1
```

Confidence Intervals

```
Variable  N      Mean
C1        44     35.09

StDev  SE Mean    95.0 % CI
11.19  1.69  ( 31.69,  38.49)
```

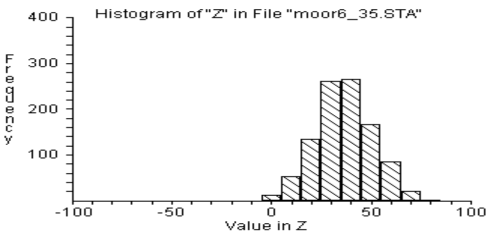
percentile z (2.5 97.5) mid95

print mid95

AVG_Z = 36.874

STAN_Z = 14.193

MID95 = 9.275   64.5




```
MTB > TwoSample 95.0 c4 c3;
SUBC>   Alternative 0.
```

Two Sample T-Test and Confidence Interval

```
Two sample T for C4 vs C3
     N   Mean    StDev SE Mean
C4  20  402.9    42.7   9.6
C3  20  366.3    50.8   11
95% CI for mu C4 - mu C3:
     ( 6.5,  67)
T-Test mu C4=mu C3 (vs not=):
T= 2.47  P=0.018  DF=  36
```

COMMENTS AND CONCLUSIONS

We have now had about two to three years of experience teaching in this mode and some insights are emerging. Our biggest fear, crashing of the hardware and/or software, never occurred because of the assistance we received from our network and lab staff who installed the software such that it could never be corrupted. Our second biggest fear, the difficulty for the students in navigating around all the software(Minitab, Resampling Stats, Word and e-mail) also never materialized; this generation of students is totally comfortable with Windows and screens.

Unfortunately, they are too comfortable. In the last year, our Department moved to a brand-new building with greatly improved computing facilities which we felt would markedly improve the hands-on aspect of the course and provide vastly better projection equipment for demonstrations. The down side of this technology is that students could hide behind the monitors to indulge in infantile e-mail and screen-saving creations instead of paying attention to the topic at hand. One of our labs does have a system(Robotel) for disabling their computers and slaving them to the instructor's; this is a feature we strongly recommend in order to avoid possible anarchy.

Although the dream we had of showing things three ways—conventionally with Minitab, simulation via Resampling Stats and simulation via Minitab—has a great deal of merit, it must be admitted that many of the students do not share this vision. For the statistically-challenged student, instead of reinforcement, confusion often reigned. At first, students had severe difficulty distinguishing Minitab's dual function as a conventional formula plugger and as a simulation device; by repeatedly emphasizing the distinction in subsequent semesters, this conflation was cleared up. However, it is a struggle to make evident that the two simulation packages are doing the same thing and only the syntax is different.

When asked which simulation was preferred, the students seem to be evenly split. Perhaps this shouldn't be surprising since the two of us differ on this as well! We have reluctantly decided to eliminate Minitab as a simulation package. It is hoped that this will result in more time available for topics such as Simpson's paradox or multiple regression, which have been squeezed out of the course.

A final comment is in order and it is an echo of a much-used phrase: there is no such thing as a magic bullet. Technology, no matter how wonderful and satisfying to an accomplished practitioner, can not be guaranteed to succeed. Analogies with medicine and mathematics are relevant. Tuberculosis is curable via antibiotics but many of the current TB victims are drug users and are unable or unwilling to follow the instructions regarding frequency and length of use of the antibiotics, rendering the efficacious drugs useless. Likewise, great hopes made for teaching calculus by means of software such as Mathematica have proven to be unrealized despite the undeniable ease the software provides for doing the dirty work of differentiating and integrating. The situation in statistics is quite similar. Drudgery can be reduced but the reality of the current educational situation is that our students are employed far too many hours while carrying a full academic load. They therefore devote too little time to the course; consequently, any technique no matter how pedagogically sensible and worthwhile can run into impediments difficult to surmount. Nonetheless, the customary, formulaic way of teaching elementary statistics has proven to be so unsuccessful that this and other innovations are worth exploring even though none is likely to be a panacea.

REFERENCES

Moore, D. S. (1995). *The Basic Practice of Statistics*, W.H. Freeman, New York.
Simon, J. (1992). *Resampling: The New Statistics*, Duxbury Press, Belmont, Ca.