

Presenting the results

22S:166 Computing in Statistics
Simulation studies in statistics, continued
Lecture 13
October 10, 2007

Based on a lecture by Marie Davidian for
ST 810A - Spring 2005
Preparation for Statistical Research
North Carolina State University
<http://www4.stat.ncsu.edu/~davidian/st810a/>

What did you do and why? Before giving results, you must first give a reader enough information to appreciate them!

- State the *objectives* – Why do this simulation? What specific questions are you trying to answer?
- State the *rationale* for choice of factors studied, assumptions made
- Review all *methods* under study – be precise and detailed
- Describe *exactly* how you generated data for each choice of factors – enough detail should be given so that a reader could write his/her *own program* to reproduce your results!

1

2

Some basic principles:

- Present only a *subset* of results (“Results were qualitatively similar for all other scenarios we tried.”)
 - i.e., don’t overwhelm the reader with a dozen tables of numbers that all carry the same message
- Present only interesting, useful information (“Relative biases for all estimators were less than 2% under all scenarios and hence are not shown in the table.”)
ritem The *mode of presentation* should be *friendly* . . .

3

Tables

Tables are an obvious way to present results; *however*, some caveats

- Avoid zillions of numbers jam-packed into a table!
- Place things to be compared adjacent to one another so that comparison is easy
- Rounding. . .

4

Rounding: Three reasons (Wainer, 1993)

- Humans cannot understand more than two digits very easily
- More than two digits can almost never be statistically justified
- We almost never care about accuracy of more than two digits

Wainer, H. (1993) Visual Revelations, *Chance Magazine*

5

Examples

- “This year’s school budget is \$27,329,681.32” *or* “This year’s school budget is about 27 million dollars”
- “Mean life expectancy of Australian males is 67.14 years” *or* “Mean life expectancy of Australian males is 67 years”

6

Statistical justification:

We are statisticians! For example

- Reporting Monte Carlo power – how many digits?
- *Design the study* to achieve the desired accuracy and report only what we can *justify as accurate*
- The program yields 0.56273
- If we wish to report 0.56 (*two digits*) need the *standard error* of this *estimated proportion* to be ≤ 0.005 so we can tell the difference between 0.56 and 0.57 or 0.58 ($1.96 \times 0.005 \approx 0.01$)
- $d = 0.005 = 1/\sqrt{4S}$ gives $S = 10000!$

Always report the standard error of entries in the table so a reader can gauge the accuracy!

7

Bad table: Digits, “apples and oranges”

	Sample mean		Trimmed mean		Median	
	Normal	t_5	Normal	t_5	Normal	t_5
Mean	0.98515	0.98304	0.98690	0.98499	0.99173	0.98474
Bias	-0.01485	-0.01696	-0.01310	-0.01501	-0.00827	-0.01526
SD	0.33088	0.33067	0.34800	0.31198	0.39763	0.35016
MSE	0.10959	0.10952	0.12116	0.09746	0.15802	0.12273
Rel. Eff.	1.00000	1.00000	0.90456	1.12370	0.69356	0.89238

8

Good table: Digits, “apples with apples”

	Normal			t_5		
	Sample mean	Trim mean	Median	Sample mean	Trim mean	Median
Mean	0.99	0.99	0.99	0.98	0.98	0.98
Bias	-0.01	-0.01	-0.01	-0.02	-0.02	-0.02
SD	0.33	0.35	0.40	0.33	0.31	0.35
MSE	0.11	0.12	0.16	0.11	0.10	0.12
Rel. Eff.	1.00	0.90	0.69	1.00	1.12	0.89