Johns Hopkins University, Dept. of Biostatistics Working Papers

8-22-2007

EFFECTIVE COMMUNICATION OF STANDARD ERRORS AND CONFIDENCE INTERVALS

Thomas A. Louis
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, tlouis@jhsph.edu

Scott L. Zeger
Department of Biostatistics, The Johns Hopkins Bloomberg School of Public Health

Effective Communication of Standard Errors and Confidence Intervals

Thomas A. Louis, PhD[1]
Scott L. Zeger, Phd

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street
Baltimore, MD 21205

Abstract

We recommend a format for communicating an estimate with its standard error or confidence interval. The format reinforces that the associated variability is an inseparable component of the estimate and it substantially improves clarity in tabular displays.

KEY WORDS: Effective Communication; Effective Display; Confidence Intervals

1. The Issues

When reporting estimates and standard errors or confidence intervals in text or tables, the standard formats, "estimate (se)" and "estimate 95% CI: (lower, upper)," do not visually connect an estimate with its uncertainty and are therefore commonly misinterpreted. A common mistake that results from the "(lower, upper)" style is to interpret the data as providing equal support to all values within the confidence interval. In addition, when estimates are tabulated with the standard error or confidence interval in the usual format, it is difficult to make comparisons across rows and down columns.

2. The remedies

To remedy these problems, we recommend displaying an estimate with its standard error using $est_{(se)}$ and presenting for and displaying an estimate with its confidence interval using the triple of percentiles, ${}_{2.5}50.0_{97.5}$. During the *JASA* editorship of the first author, these display were encouraged. Recently, Zeger and Johnson (2007) reported estimates of excess deaths in Iraq using these formats, extending the triple of percentiles to a five-number summary, ${}_{2.5}{}_{25.0}50.0_{75.0}{}_{97.5}$ (see Table 1) noting that the decreasing point size on each side of the estimate highlights decreasing likelihood. This display is reminiscent of the five-number summary e introduced by Tukey (1977).

3. Examples

To see the effectiveness of our recommendation for in-text reporting, compare the clarity and message

[1]To whom correspondence should be addressed, tlouis@jhsph.edu

of, "the estimate is 1.48 $(se = .09)$ to $1.48_{(.09)}$ and the clarity and message of "we estimate excess deaths to be 654 $(95\%CI : 393$ to $943)$." to that of, We estimate excess deaths to be ${}_{393}654_{943}$." The recommended format is easier to read and communicates the message that the uncertainty measure IS part of the estimate.

Tabulations using the new methods pay even bigger dividends in clarity. Note the ease of making row and column comparisons using Table 2 rather than Table 3. Similarly, Tables 4 and 5 are far more effective than a table using the usual format.

4. Discussion

The proposed displays that connect a point estimate with its standard error or confidence interval in one "diagram," using different subscripting and differentially-sized fonts has several advantages. It makes clear that the estimate comprises the point estimate AND its uncertainty expressed as graphical unit. The use of subscripting and different font size substantially enhances the readers ability to compare estimates across rows or columns without the measures of uncertainty getting in the way (and vice versa). With confidence limits displayed before and after the point estimate in smaller font, it gives the correct impression that the data are more consistent with the point estimate than with the limits of the interval. The five number summary graphic is reminiscent of a likelihood function or posterior distribution. The proposed displays are easy to implement (see the LaTex code in the Appendix).

In summary, statisticians can usefully invest more effort in communicating both point estimates and measures of uncertainty. We recommend graphical displays that treat the estimate and its uncertainty as a single unit that can be combined with others in effective summaries of a statistical analyses.

5. Appendix: LaTex Code

```
\newcommand{\estse}[2]{{#1}_{(#2)}}

\newcommand{\cithree}[3]{_{{#1}}{#2}_{{#3}}}

\newcommand{\cifive}[5]{_{_{#1}{#2}}{#3}_{#4_{#5}}} .
```

Bibliography

1. Barnard, J. , Frangakis, C. E., Hill, J. L., and Rubin, D. B. (2003), "Principal Stratification Approach to Broken Randomized Experiments: A Case Study of School Choice Vouchers in New York City (with discussion)," *J. Am. Statist. Assoc.,* **98:** 299-323.

2. Hoeting, J. A., Tweedie, R. L., and Olver, C. S. (2003), "Transform Estimation of Parameters for Stage-Frequency Data," *J. Am. Statist. Assoc.,* **98:** 503-514.

3. Tukey, J. W. (1977), *Exploratory Data Analysis.* Addison-Wesley, Reading, Massachusetts.

4. Zeger, S. L., and Johnson, E. (2007), "Estimating excess deaths in Iraq since the US–British-led invasion," *Significance*, **vol. 4, #2:** 54-59.

654 (393 to 943)
$_{393}654_{943}$
$_{393}{}_{560}654_{748}{}_{943}$

Table 1: Figure 2 from Zeger and Johnson, 2007: "Alternative graphical presentations of confidence or posterior intervals: top, standard 95% interval; middle, graphical presentation of the same, demonstrating the higher likelihood or greater posterior probability in the middle of the interval; bottom, five-number summary with best estimate, 50% and 95% intervals."

	Stage		
Method	1	2	3
True	1.5	1.5	1.5
MLE	$1.48_{(.09)}$	$1.50_{(.09)}$	$1.40_{(.08)}$
MOM	$1.38_{(.17)}$	$1.35_{(.23)}$	$1.69_{(.21)}$
LT (iterated)	$1.37_{(.19)}$	$1.29_{(.25)}$	$1.57_{(.42)}$
LT (minimum variance)	$1.37_{(.16)}$	$1.32_{(.22)}$	$1.54_{(.34)}$

NOTE: Estimated standard errors given in parentheses.

Table 2: Table 1 from Hoeting et al., 2003 (as published):
"Parameter estimates for simulated data."

	Stage		
Method	1	2	3
True	1.5	1.5	1.5
MLE	1.48 (.09)	1.50 (.09)	1.40 (.08)
MOM	1.38 (.17)	1.35 (.23)	1.69 (.21)
LT (iterated)	1.37 (.19)	1.29 (.25)	1.57 (.42)
LT (minimum variance)	1.37 (.16)	1.32 (.22)	1.54 (.34)

NOTE: Estimated standard errors given in parentheses.

Table 3: Table 1 from Hoeting et al., 2003 (converted to the usual format):
"Parameter estimates for simulated data."

Grade at application	Applicant's school: Low		Applicant's school: High	
	Reading	Math	Reading	Math
1	$3.4_{(-2.0,8.7)}$	$7.7_{(3.0,12.4)}$	$1.9_{(-7.3,10.3)}$	$7.4_{(.2,14.6)}$
2	$.7_{(-3.7,5.0)}$	$1.9_{(-2.4,6.2)}$	$-.9_{(-9.4,7.3)}$	$1.5_{(-6.2,9.3)}$
3	$1.0_{(-4.1,6.1)}$	$5.0_{(-.8,10.7)}$	$-.8_{(-9.5,7.7)}$	$4.0_{(-4.9,12.5)}$
4	$4.2_{(-1.5,10.1)}$	$4.3_{(-1.6,10.1)}$	$2.7_{(-6.3,11.3)}$	$3.5_{(-4.7,11.9)}$
Overall	$2.2_{(-.9,5.3)}$	$4.7_{(1.4,7.9)}$	$.6_{(-7.1,7.7)}$	$4.2_{(-2.6,10.9)}$

NOTE: Plain numbers are means, and parentheses are central 95% intervals of the posterior distribution of the effects on percentile rank.

Table 4: Table 6 from Barnard et al., 2003 (as published): "ITT Effect of Private School Attendance on Test Scores."

Grade at application	Applicant's school: Low		Applicant's school: High	
	Reading	Math	Reading	Math
1	$_{-2.0}3.4_{8.7}$	$_{3.0}7.7_{12.4}$	$_{-7.3}1.9_{10.3}$	$_{0.2}7.4_{14.6}$
2	$_{-3.7}0.7_{5.0}$	$_{-2.4}1.9_{6.2}$	$_{-9.4}-0.9_{7.3}$	$_{-6.2}1.5_{9.3}$
3	$_{-4.1}1.0_{6.1}$	$_{-.8}5.0_{10.7}$	$_{-9.5}-0.8_{7.7}$	$_{-4.9}4.0_{12.5}$
4	$_{-1.5}4.2_{10.1}$	$_{-1.6}4.3_{10.1}$	$_{-6.3}2.7_{11.3}$	$_{-4.7}3.5_{11.9}$
Overall	$_{-0.9}2.2_{5.3}$	$_{1.4}4.7_{7.9}$	$_{-7.1}.6_{7.7}$	$_{-2.6}4.2_{10.9}$

NOTE: Plain numbers are means, and parentheses are central 95% intervals of the posterior distribution of the effects on percentile rank.

Table 5: Table 6 from Barnard et al., 2003 (converted to recommended format): "ITT Effect of Private School Attendance on Test Scores."