DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 16 | Issue 1 Article 42

5-1-2017

Errors in a Program for Approximating Confidence Intervals

Andrew V. Frane
University of California Los Angeles, avfrane@gmail.com

Recommended Citation
Frane, A. V. (2017). Errors in a program for approximating confidence intervals (Letter to the Editor). Journal of Modern Applied Statistical Methods, 16(1), 779-782. doi: 10.22237/jmasm/1493599320

Journal of Modern Applied Statistical Methods
May 2017, Vol. 16, No. 1, 779-782.
doi: 10.22237/jmasm/1493599320

ISSN 1538 − 9472

Errors in a Program for Approximating Confidence Intervals

Andrew V. Frane
University of California Los Angeles
Los Angeles, CA

An SPSS script previously presented in this journal contained nontrivial flaws. The script should not be used as written. A call is renewed for validation of new software.

Keywords: effect size, confidence intervals, SPSS, syntax

Letter to the Editor

Walker (2015) presented an SPSS program for estimating effect sizes and approximating confidence intervals. It contains flaws and should not be used. The consequences are nontrivial, as is apparent from Walker's example, which used the following input: $M_1 = 16.45$, $M_2 = 11.77$, $SD_1 = 2.23$, $SD_2 = 4.66$, $N_1 = 30$, $N_2 = 34$, $N = 64$, where M_1 and M_2 are the sample means, SD_1 and SD_2 are the sample standard deviations, N_1 and N_2 are the group sample sizes, and N is the total sample size. Given this input, the resulting 95% confidence intervals in Walker's output (see his Table 1) are far too narrow: either [1.109, 1.403] or [1.094, 1.387], depending on whether Cohen's *d* or an approximation of Hedges' *g* is used in the estimation.

Walker did not validate these results by simulation, or by analytic methods, or by comparing the results to those produced by established software. For example, the `ci.smd` function in the extensively vetted MBESS package for *R* (see Kelley, 2007; Kelley & Rausch, 2006) uses a standard iterative procedure to compute exact confidence intervals for the standardized effect size. For Walker's input, the `ci.smd` function may be executed in conjunction with the `smd` function, as follows:

```
library (MBESS)
```

Mr. Frane is a doctoral student of cognitive psychology in the Multisensory Perception Lab. Email at avfrane@gmail.com.

```
cohend <- smd (Mean.1=16.45, Mean.2=11.77, s.1=2.23, s.2=4.66,
   n.1=30, n.2=34)
ci.smd (smd=cohend, n.1=30, n.2=34, conf.level=.95)
```

This method correctly gives the 95% confidence interval as [0.714, 1.790]. Note that this interval is much wider than Walker's approximations and is appropriately asymmetrical around Cohen's *d*.

Part of the problem with Walker's code is how it computes the variables it calls D1 and G1. These cryptically-named variables purportedly estimate the error terms of Cohen's *d* and Hedges' *g* (respectively), but as coded actually estimate the squares of those error terms. That is, the program computes estimated variances when it should be computing estimated standard errors. The same confusion is evident in Walker's equation 9 (compare to Hedges & Olkin, 1985, p. 86, equation 15, which appropriately squares the error term on the left side of the equation). Hence, Walker's erroneous computations could be vastly improved by adding square roots to the two lines of code where D1 and G1 are computed, as follows:

```
COMPUTE D1 = SQRT (N / (N1*N2) + COHEND**2 / (2*N)).
COMPUTE G1 = SQRT (N / (N1*N2) + HEDGESG**2 / (2*N)).
```

However, there is no justification for using approximations at all, given that superior, exact confidence intervals can now be easily computed with simple commands in freely available, industry standard software (namely, *R* with the MBESS package).

Walker acknowledged that by disregarding noncentrality, the program could not provide exact confidence intervals, a limitation defended as follows: "Bird (2002) found that if *d* is < 2.00, which in social science research frequently can be the circumstance with middling-sized effects (Richard, Bond, & Stokes-Zoota, 2003; Rosnow & Rosenthal, 2003), adjustment for noncentrality is not compulsory" (Walker, 2015, p. 285). Bird (2002) did note that heuristically speaking, approximate standardized intervals are likely to be similar to exact standardized intervals for $d < 2$, provided degrees of freedom ≥ 30. However, Walker overlooked Bird's caveat that "exact standardized intervals should be preferred to approximate standardized intervals whenever both are available" (Bird, 2002, p. 204).

Walker's program implements incorrectly a method that would be obsolete even if implemented correctly. The program also contains other peculiarities. For example, given that the user must input N_1 and N_2, it is redundant that the program also requires the user to input N (which the program could instead have computed for itself, as simply $N_1 + N_2$). Additionally, an anonymous reviewer of the present letter identified a potentially confusing conflict between the coding and the text in Walker's article: The coding computes Cohen's *d* using the pooled standard deviation, which is likely the proper approach, but Walker's equation 6 computes Cohen's *d* using the unweighted average of SD_1 and SD_2.

Walker (2015) appeared in the same issue as an article noting the perils of using inadequately vetted statistical software (Lorenz, Markman, & Sawilowsky, 2015). Indeed, checking new software against established software prior to dissemination and professional use is essential.

References

Bird, K. D. (2002). Confidence intervals for effect sizes in analysis of variance. *Educational and Psychological Measurement, 62*(2), 197-226. doi: 10.1177/0013164402062002001

Hedges, L. N., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Kelley, K. (2007). Confidence intervals for standardized effect sizes: Theory, application, and implementation. *Journal of Statistical Software, 20*(8), 1-24. doi: 10.18637/jss.v020.i08

Kelley, K., & Rausch, J. R. (2006). Sample size planning for the standardized mean difference: Accuracy in parameter estimation via narrow confidence intervals. *Psychological Methods, 11*(4), 363-385. doi: 10.1037/1082-989X.11.4.363

Lorenz, A. J., Markman, B. S., & Sawilowsky, S. (2015). Caution for software use of new statistical methods (R). *Journal of Modern Applied Statistical Methods, 14*(2), 275-281. Retrieved from http://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=2021&context=jmasm

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7*(4), 331-363. doi: 10.1037/1089-2680.7.4.331

Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, *57*(3), 221-237. doi: 10.1037/h0087427

Walker, D. A. (2015). Two group program for Cohen's *d*, Hedges' *g*, η^2, R_{adj}^2, ω^2, ε^2, confidence intervals, and power. *Journal of Modern Applied Statistical Methods*, *14*(2), 282-292. Retrieved from http://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=2086&context=jmasm