Letter to the Editor

From

Martin R. Petersen[*,1] and James A. Deddens[2]

[1] NIOSH MS-R15, 4676 Columbia Parkway, Cincinnati OH 45226-1998, USA
[2] NIOSH MS-R15, 4676 Columbia Parkway, Cincinnati OH 45226-1998, USA
  and, Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45215

Regarding Paper

"Parameter Estimation and Goodness-of-Fit in Log Binomial Regression"

by Blizzard L. and Hosmer D.W.

Biometrical Journal (2006) **48**, 5–22

Blizzard and Hosmer (2006) extend the use of diagnostics from logistic models to log binomial models, which is a very useful addition to the literature. Our comments are somewhat peripheral to this main topic.

First, we agree that one should be concerned when estimated probabilities exceed unity (i.e. an inadmissible estimate). In theory, this can be avoided by using maximum likelihood estimation, which has other useful properties as well and is generally considered the method of choice. In practice, however, software fails to converge when the maximum likelihood estimated probability is exactly one, which often happens when an independent variable is continuous. Deddens, Petersen, and Lei (2003) successfully solved this problem in theory and, at least for SAS, in practice. The method, called the COPY method due to the nature in which the data set is manipulated, simply approximates the maximum likelihood estimation by making the estimated probabilities slightly less than one at the offending points. Blizzard and Hosmer did not use the method because they still had convergence problems in Stata even though no estimated probabilities were exactly one. Barros and Hirakata (2003) first proposed the use of the Poisson model with a robust variance estimator, and they also had convergence problems with the COPY method in Stata. In a review of this paper prior to publication, Deddens (2003) suggested that the problem may have been due to the chosen start values. Barros and Hirakata (2003) subsequently reported that the use of the "search" option solved their convergence problems in Stata. We do not use Stata, but perhaps Blizzard and Hosmer would be interested in checking to see if this option will solve their Stata convergence problems.

Second, as the COPY method was proposed by Deddens et al. (2003), the data were only manipulated when the log binomial model failed to converge. Thus Blizzard and Hosmer have performed the COPY method for the 1000 simulations that they used. They simply discarded the simulations for which data manipulation was required. In our experience, it is on these simulations that the inferiority of the Poisson approach generally becomes more apparent, and inadmissible estimates become more common. If Blizzard and Hosmer can get the COPY method to converge in Stata, their diagnostics can be used for maximum likelihood based estimation on any data set without having to switch from Stata to SAS.

---

 * Corresponding author: e-mail: mrp1@cdc.gov

References

Barros, A. and Hirakata, V. (2003). Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Medical Research Methodology* **3**, 21. http://www.biomedcentral.com/1471-2288/3/21.

Blizzard, L. and Hosmer, D. W. (2006). Parameter estimation and goodness-of-fit in log binomial regression. *Biometrical Journal* **48**, 5–22.

Deddens, J. A. (2003). RE: Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio, *BMC Medical Research Methodology* **3**, 21. http://www.biomedcentral.com/1471-2288/3/21/prepub.

Deddens, J. A., Petersen, M. R., and Lei, X. (2003). Estimation of prevalence ratios when PROC GENMOD does not converge. *Proceedings of the 28th Annual SAS Users Group International Conference. March 30–April 2, 2003*. http://www2.sas.com/proceedings/sugi28/270-28.pdf.

Petersen, M. R. and Deddens, J. A. (2006). Letter to the editor: RE: Easy SAS calculations for risk or prevalence ratios and differences. *American Journal of Epidemiology* **163**, 1157–1163.