Computational Statistics & Data Analysis 2 (1984) 79–80
North-Holland

Rejoinder

Morton B. BROWN
Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA

Camil FUCHS
Department of Statistics, Tel-Aviv University, Tel-Aviv 69978, Israel

Received March 1984

In [1] we describe how reading a dataset into a computer program in the usual cases-by-variables structure can implicitly cause errors when a log-linear model is fitted to the data using a Newton–Raphson algorithm. Our example treats problems that arise when cells with zero frequencies are excluded from the dataset Aston and Wilson in their comment show that when all cells appear in the dataset (including cells with zero expectations), there can be errors produced by the Newton–Raphson algorithm (such as in the degrees of freedom and the parameter estimates).

Aston and Wilson claim to "outline a simple and straightforward method for overcoming such problems". Their method is "to identify the occurrence of zeros in any of the marginal configurations defined by the particular log-linear model and constrain the corresponding (zero) cells to have estimated cell frequencies of zero". That approach coincides with our recommendation that "when using the NR algorithm, the vector of frequencies used as input must include all cells but those having zero expected values under the model to be fitted" [1, p. 12]. Constraining a cell to be zero is equivalent to omitting the cell from the input vector.

Aston and Wilson do not indicate any novel method of finding and eliminating the possible zeros. On the contrary they require a two-pass procedure: after the first pass the existence of aliasing is identified and in the second pass the cells corresponding to marginal zero are constrained to zero.

Aston and Wilson then compared the results of a log-linear analysis with that of a logistic analysis of the same data. They observed that the parameters estimates and their standard errors were identical. When there is aliasing of parameter estimates, the choice of parameters that are estimated can be arbitrary and vary from program to program. The fact that in this case Aston and Williams fortunately obtained the same parameter estimates does not negate this fact.

The comparison of models M_1 and M_2 in [1] states that "when the number of estimable parameters is less than the number of parameters, it is not possible to replace the model by the model excluding the nonestimable parameters". The same is correct for the comparison of models M_3 and M_4. Aston and Wilson do not contradict this but add that "this is because the parameter spaces are not nested". In [1] there is no suggestion that the difference between the models is asymptotically chi-square. Aston and Wilson recommend a simulation study but do not describe it in adequate detail. (Their aside about typographical errors in M_2 refers to the omission of two terms, DE and DB, from the model. The numerical results are for the correct model. A second typo is in Table 1 where the correct frequency for cell 21212 is 1, and not 2 as reported in the table.)

The comments of Aston and Wilson do not invalidate any of the conclusions in [1]. It strengthens our opinion that a researcher needs to be aware of the problems that can arise with sparse data and recognize the pitfalls in both the Newton–Raphson and the iterative proportional fitting algorithms.

[1] M.B. Brown and C. Fuchs, On maximum likelihood estimation in sparse contingency tables, *Computational Statistics & Data Analysis* **1** (1983) 3–15.