Rejoinder to Kullback and Keegel note on the minimum discrimination information approach

Morton B. BROWN

Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA

Kullback and Keegel [3] (K & K) espouse the minimum discriminant information (MDI) approach. For the problem considered by Brown and Fuchs [1] (B & F), K & K agree that "the results of the MDI procedure are the same as the maximum likelihood estimates and the MDI statistic... is the log-likelihood ratio statistic". Therefore, the two methods must yield identical results when the same model is being fitted to the same set of data. One method cannot be more correct than the other.

K & K then demonstrate that by using their program, they obtain the same results using a Newton–Raphson (NR) algorithm as by using an iterative proportional fitting (IPF) algorithm (see Table 1). In B & F different results were obtained by the two different algorithms. The difference in B & F occurred because the data sets to which the model was being fitted differed between the two algorithms. K & K forced the same dataset to be used by both algorithms.

A major conceptual issue which arises in several of the specific comments of K & K is whether the definition of the model also includes the definition of the observation space. B & F used the term model to refer to the generic form of the model which can be defined even before data are collected. This generic form is specified to a computer program which then reads the data matrix presented to it and computes the parameter estimates of the model. When the data presented are changed, the estimates change even if the generic model is the same. In B & F we considered that the data presented to the Newton–Raphson algorithm were that specified in our Table 1 (which excluded all the observed zeros) whereas the program using an IPF algorithm first completes the frequency table and in that process modifies the set of observations. Therefore, the two algorithms used by B & F were fitting the same generic model to two sets of observations which differed in their zero frequencies.

On the other hand K & K assume that there is a dependent variable with two levels and add a zero frequency whenever one and only one of the levels is present

(see Table 1), but omit pairs with two zeros. Then they replace these added zeros by a small positive constant (0.000001). If they had omitted all zeros, they would have obtained the results for NR reported in B&F; if they had included all zeros, they would have obtained incorrect degrees of freedom. Similarly, B&F was able to obtain the correct chi-square using the Newton–Raphson algorithm by choosing the appropriate set of observed cells.

With respect to the specific comments:

(1) The choice of linearly independent constraints is a function of the set of cells to be included in the model. K&K chose to eliminate certain cells a priori. Otherwise, they may also have a problem with the rank of the constraints.

(2) This is a matter of parametrization and can be set in any manner desired.

(3, 7) These corrections were noted in a previous rejoinder [2].

(4) L_1 includes the extra zeros since it looks at pairs of cells. If one of a pair is not present, its frequency is set to zero.

(5) M_1 is a model defined in a generic sense. The observation space to be used for the model is defined by the data read as input.

(6) The difference in DF I believe was due to the tolerance for inversion. Due to near singularity the rank was computed incorrectly by the program.

The MDI approach and the maximum likelihood approach produce the same results for the problem described by B&F. The choice of which to use is a matter of convenience by the practitioner. The lack of reference to MDI by B&F was because discussion of the MDI approach did not add to the numerical issues being described.

References

[1] M.B. Brown and C. Fuchs, On maximum likelihood estimation in sparse contingency tables, *Comput. Statist. Data Anal.* 1 (1983) 3–15.
[2] M.B. Brown and C. Fuchs, Rejoinder, *Comput. Statist. Data Anal.* 2 (1984) 79–80.
[3] S. Kullback and J.C. Keegel, The minimum discrimination information approach to the analysis of contingency tables, *Comput. Statist. Data Anal.* 3 (1985) 133–141 (this issue).