Rejoinder

Alan Agresti

I thank the discussants for their comments. I appreciate their compliments, as well as their criticisms and suggestions, which nicely supplement the presentation in my paper. It is reassuring to see that, despite what some might call the excessive length of my paper, alternative perspectives raise yet other noteworthy issues and cast additional light on this subject.

In my article, I mentioned a need for the development of exact methods for model checking, and I am pleased to see the contribution by Ed Bedrick and Joe Hill on this topic. They suggest a simple algorithm for generating the relevant reference set for logistic regression. Their algorithm would seem to generalize to loglinear models.

Regarding the problem noted by Bedrick and Hill of potential near degeneracy in using conditional methods, I am afraid I do not see any simple solutions (other than perhaps becoming a Bayesian). Degeneracy is the most extreme form of the severe discreteness that can occur with conditional methods. The severe discreteness is the primary weakness of this type of method and is, I believe, at the heart of the objection many statisticians have with methods such as Fisher's exact test. An approximate solution in logistic regression is to slightly collapse the data in order to produce a fuller conditional distribution of the data.

I thank Bedrick and Hill for clarifying and extending my remark about exact analysis for parameters when the link function in a generalized linear model is noncanonical. They show that some types of conditional inference may still be useful, both for model checking and inference about a parameter of interest. I would like to see them develop this discussion further for some interesting noncanonical models for categorical data. Likewise, I would like to see further discussion of their view treating Fisher's exact test only as a goodness-of-fit test. This is subtly distinct from the usual view of its also serving as a test comparing two independent binomials. Perhaps they can help to clarify this longstanding controversy, although I do not expect to see statisticians reach agreement about how to analyze 2×2 tables, at least not in my lifetime.

I think that both of Diane Duffy's ideas merit considerable attention. We statisticians commonly complain that users of statistics pay too much attention to p-values and statistical tests at the expense of more informative types of analysis. Perhaps we can at least convince them to perform sensitivity analyses, such as the ones Duffy recommends, so that they do not take their p-values too literally, feeling compelled to report them to several decimal places.

When $n_{1+} = n_{2+}$, deleting an observation from one row has the same impact on the magnitude of the one-sided p-value (but in the opposite direction) as deleting the same type of observation from the other row. Thus, in this balanced case, the observed p-value falls in the middle of the interval (P_L, P_U) for Duffy's first type of perturbation. For instance, for table $(10, 90/20, 80)$, the one-sided p-value is 0.0367, with $P_L = 0.0231$ and $P_U = 0.0503$. As shown by her example, this need not happen for the usual two-sided p-value, for which deletion of an observation sometimes has no effect on the p-value.

Duffy suggests studying whether algorithms for exact analyses can be adapted to aid Bayesian computations. She also proposes a conditional Bayesian analysis and suggests comparing it to ordinary Bayesian and frequentist procedures. The recent significant improvements in computational tools (e.g., Gibbs sampling) for Bayesian methods suggests that we should soon see Bayesian and hybrid methods more fully developed for multidimensional contingency tables. For sparse contingency tables, several problems exist for which unconditional frequentist approaches fail and for which a Bayesian approach would be a natural alternative to a conditional approach. An example is the analysis of $2 \times 2 \times K$ tables with large K. When the true odds ratio is identical in each 2×2 table, its unconditional ML estimator is inconsistent when K grows at the same rate as n, such as when each table consists of a case-control matched pair (e.g., Breslow and Day, 1980, page 250). When the true odds ratios are not identical in each table, a Bayesian or empirical Bayesian approach would seem to be a reasonable way of smoothing the stratum-specific estimators of odds ratios, borrowing from the whole to get estimators with improved MSE properties.

Leonardo Epstein and Stephen Fienberg also suggest the worthiness of a Bayesian perspective, pointing out that it may be no more difficult computationally than exact conditional methods. It is interesting to note that, before much development of methodology for multiway contingency tables had taken place, Lindley (1964) argued that Bayesian methods had the advantage (compared to