tingency tables are: (i) jackknife-type perturbations that decrease each count by one in turn, (ii) perturbations that involve increasing or decreasing each count by one in turn, (iii) perturbations loosely based on a notion of misclassification that preserves the total sample size but reallocates up to a certain fraction of the observations, and (iv) more general perturbations that need not preserve total sample size and also permit more than one cell count to alter. Schemes such as (i) and (ii) have a certain natural appeal in moderate-to-large size contingency tables—one would like to think that changing just one cell by one count could not seriously effect the p-value. In a context where protection against misclassification is desired, a scheme of type (iii) is appropriate. A misclassification-based scheme may differ depending on what, if any, margins are fixed by the sampling design. For example, in a 2×2 table with fixed column margin, one may be primarily concerned with potential errors in row classification. In other words, one could want a perturbation scheme that preserved the column margin.

Working with a set of approximately one dozen real examples of 2×2 tables culled from assorted textbooks, the effects of perturbation schemes of types (i), (ii) and (iii) on Fisher's exact test were studied. Denote the actual p-value by P and the minimum and maximum p-values achieved over the set of perturbations by P_L and P_U, respectively. The following tentative conclusions rest on this limited experience; in the interests of space, I will illustrate the points exclusively with respect to the table given by Agresti in Section 2.1 having counts by row $(10, 90/20, 80)$. Scheme (ii) may be preferable to scheme (i), as there are cases where scheme (i) alters the p-value in only one direction so that one of P_L or P_U equals P, while scheme (ii) has $P_L \neq P$ and $P_U \neq P$. For $(10, 90/20, 80)$, $P = 0.073$; under scheme (i), $P_L = 0.043$ and $P_U = P = 0.073$; under scheme (ii), $P_L = 0.043$ and $P_U = 0.082$. Under scheme (ii), it is frequently, but by no means always, the case that P_L and P_U are achieved by increasing and decreasing the same cell. For $(10, 90/20, 80)$, this is the case with P_L arising from the table $(9, 90/20, 80)$ and P_U arising from the table $(11, 90/20, 80)$. Scheme (iii) often leads to a much wider range of p-values. For (10,90/20,80), moving one count under scheme (iii), $P_L = 0.028$ for $(9, 90/21, 80)$ and $P_U = 0.117$ for $(11, 89/20, 80)$. In the event that scheme (iii) is restricted to perturbations that preserve the row margin, $P_L = 0.043$ for $(9, 91/20, 80)$ and $P_U = 0.117$ is unchanged.

It is possible that further work on the sensitivity of exact inference may lead to rough guidelines on, say, the percentage change in p-value corresponding to some set of perturbations for the data (see Dupont, 1986). For the present time, however, it would be helpful if software packages were setup to easily permit sensitivity analysis based on these or other perturbation schemes. Developing techniques to permit efficient sensitivity analyses (i.e., without repeating the computation for each perturbed table) would be a useful area for research.

Comment

Leonardo D. Epstein and Stephen E. Fienberg

We would like to congratulate Professor Agresti for his thorough review of the recent literature on exact inference in contingency tables and for organizing it in a way that allows us to focus on some key statistical issues. Our first observation relates to the work "exact," which has an everyday meaning that may not coincide with its technical meaning in the present context. It is also a value-laden descriptor that suggests that any statistical method that is not exact may not be very good. As the following comments imply, nothing could be further from the truth.

The most widely studied problem involving categorical data, and seemingly the simplest, is that of drawing inferences for the risk ratio and risk difference in 2×2 contingency tables. Yet, this simple situation highlights many of the most controversial aspects of statistical methodology and theory. Before discussing these issues, we note that there are few practical statistical problems that come in the simple form of a 2×2 table. Most investigations involve a large number of variables, both continu-

Leonardo D. Epstein is an Assistant Professor in the Department of Biostatistics at the Johns Hopkins University, 615 North Wolfe Street, Baltimore, Maryland 21205. Stephen E. Fienberg is a Professor of Statistics and Law, and Vice President of Academic Affairs at York University, 4700 Keele Street, North York, Ontario M3J 1P3, Canada.