GOOD, I. J. (1950). *Probability and the Weighing of Evidence*. Hafner, New York.
GOOD, I. J. (1976). On the application of symmetric Dirichlet distributions and their mixtures to contingency tables. *Ann. Statist.* **4** 1159–1189.
GOOD, I. J. and CROOK, J. F. (1980). The information in the margins of a 2 × 2 contingency table. Unpublished report.
GÛNEL, E. (1982). Bayes factors in three-way contingency tables. *Comm. Statist. A—Theory Methods*. **11** 911–931.
GÛNEL, E. and DICKEY, J. (1974). Bayes factors for independence in contingency tables. *Biometrika* **61** 545–557.
JEFFREYS, H. (1939). *Theory of Probability* (3rd ed. 1961) Clarendon Press, Oxford.
KADANE, J. B. and DICKEY, J. M. (1980). Bayesian decision theory and the simplification of models. In *Evaluation of Econometric Models*. (J. Kmenta and J. Ramsey, eds.) pp. 245–268. Academic, New York.
LI, L.-A. (1983). Decomposition theorems, conditional probability, and finite mixture distributions. Ph.D. dissertation, Dept. of Mathematics and Statistics, SUNY at Albany.
LINDLEY, D. V. (1971). *Bayesian Statistics, A Review*. SIAM, Philadelphia.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
STATE UNIVERSITY OF NEW YORK AT ALBANY
ALBANY, NEW YORK 12222

STEPHEN E. FIENBERG[1]

Carnegie-Mellon University

Diaconis and Efron's (henceforth DE) goal in this paper is a laudable one—to help interpret the classical chi-square statistic used to test for independence in two-way contingency tables in cases where independence clearly does not hold. Their mathematical statistics results are impressive, their theorems are seemingly impeccable, and their writing style is lucid. Yet, even after several readings, I came away from the paper with a feeling of disquiet, and a belief that they had failed to achieve their goal for most practical purposes. This comment provides some explanations for my disquiet and raises questions about the immediate utility of DE's results. The claim here is not so much that DE's results will not be of use to someone in the future (for their elegant results and geometrical interpretations will surely be put to good use), but rather that they will not be useful for the purpose originally proposed.

The statistical model for the counts in a two-way contingency table has two components: (1) a sampling model for the generation of the counts given a set of cell probabilities or expected values; (2) a structural model (corresponding to a curved manifold in the simplex) for the cell probabilities that is typically tied to the relationship between the categorical variables underlying the rows and

[1] Research partially supported by Office of Naval Research Contract N00014-84-K-0588, by a fellowship from the John Simon Guggenheim Foundation, and by NSF Grant BNS-8011494 while the author was a fellow at the Center for Advanced Study in the Behavioral Sciences.

columns (e.g., see Bishop, Fienberg and Holland, 1975, for an elaboration of these two components). As lucid as their style is, DE gloss over the distinction between these components, and, in doing so, they lead the unwary reader down the curve of a one-dimensional exponential family of distributions to the limiting case of the "uniform" distribution and the volume test. DE's original null hypothesis, H_1, assumes a multinomial sampling model *and* the structural model of independence. The usual alternative hypothesis, H_A, changes the structural model but not the sampling model. Sometimes this alternative model for the cell probabilities is unrestricted, as in the classical chi-square test for independence, and other times it is an alternative curved manifold of the same or higher dimension, e.g., Goodman's (1979) uniform association model provides a class of alternatives to independence, indexed by a single parameter, for cross-classifications having ordered categories. DE's original alternative hypothesis, H_0, in effect changes both the sampling model and the structural model. The volume test then computes the probability content, under H_0, of the rejection region of the chi-square test (see Fienberg and Gilbert, 1970).

I am not sure that the volume test or DE's one-parameter exponential family have any direct practical interest for contingency table analysis. In most substantive problems, we know what the sampling model is. When we take a simple random sample from a population, the sampling model is multinomial for infinite populations (and approximately so for finite populations). It is hard to envision under exactly what circumstances the sampling model of H_0 would be of interest. The entire DE approach involves first the use of the chi-square test, then the volume test, and finally the estimation of θ, or equivalently a reduced sample size. I would be somewhat happier with this package were I able to identify and interpret the sampling and structural model components for specific values of θ, in cases other than $\theta = 0$. Perhaps DE have some insights into this aspect of their family of alternative models.

The interpretation of the volume test does not have a truly Bayesian flavor, despite what one might infer from a reading of Section 2 of DE. They begin by giving the vector of cell probabilities, $\boldsymbol{\pi}$, a $D_{IJ}(\mathbf{1}_{IJ})$ distribution and then they note that the marginal distribution for the vector of observed proportions, $\mathbf{p}$, is uniform over the lattice points of the simplex. But it seems to me that if we are to put on our Bayesian boots (habiliment I am more than happy to wear) we need to put them onto both feet. This means that we must come to grips with priors for the cell probability under H_1, e.g., by putting independent priors (possibly uniform) on the margins of $\boldsymbol{\pi}$. Then I guess we could get a marginal distribution for $\mathbf{p}$. Of course, we might begin by putting a nonuniform Dirichlet prior on $\boldsymbol{\pi}$, and then integrate over $\boldsymbol{\pi}$ to get a new sampling distribution for $\mathbf{p}$, e.g., as Brier (1980) does to represent cluster sampling. But then the structural model needs to be imposed on the probability vector parameter of the Dirichlet prior, e.g., independence for H_1 and unrestricted for H_A.

Preferable to DE's non-Bayesian approach, in my view, is the calculation of the Bayes factor (see Good, 1976). For example, with uniform priors on the margins of $\boldsymbol{\pi}$ under independence and a hypergeometric sampling distribution (which conditions on the sample margins, i.e., of $\mathbf{p}$), the Bayes factor is

$$\frac{n! \prod m_{ij}!}{\prod m_{i+}! \prod m_{+j}!} \times \frac{1}{N^{(n)}(\mathbf{r}, \mathbf{c})}.$$

Perhaps I am missing something, but I do not see a clear link to the volume test here. Good discusses several *calculations* considered by DE, but they are used for different purposes in his paper. The seemingly Bayesian random effects approach of Section 4 of DE is based on analogy and, in the end, I failed to find a Bayesian interpretation to go with the one-parameter exponential family of alternative in Section 5 (except in the limit when $\theta = 1$).

Despite the fact that I have not yet been won over to DE's program for the analysis of two-way contingency tables, I do believe that elegant results, such as those described in their paper, will inspire unimagined uses. Thus I would like to end this comment with a set of queries. More often than not, the statistician working on a substantive categorical data problem will be faced with several variables, not just two. Have the authors given any thought to the extensions of their approach to chi-square tests for log-linear and logit models in multi-way contingency tables? As the dimension of the table increases, the number of simple null models involving some form of complete and/or conditional independence grows extremely rapidly (e.g., see the discussion in Bishop, Fienberg and Holland, 1975; or Good, 1976). Would the authors choose to begin with the generalization of the volume test for multi-way tables, and if so, how would they index the exponential families of alternative distributions? Would they have a different exponential family for each null model? Would there be a heuristic Bayesian interpretation for the indexing parameters? Answers to these questions not only may be of interest in their own right, but they may also shed some light on the utility of DE's approach to the interpretation of the chi-square statistic in the analysis of two-way tables.

REFERENCES

BISHOP, Y. M. M., FIENBERG, S. E., and HOLLAND, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice.* M.I.T. Press, Cambridge, MA.

BRIER, S. S. (1980). Analysis of contingency tables under cluster sampling. *Biometrika* **67**, 591–596.

DIACONIS, P. and EFRON, B. (1985). Testing for independence in a two-way table: New interpretations of the chi-square statistic. *Ann. Statist.* **13** 845–874.

FIENBERG, S. E. and GILBERT, J. P. (1970). The geometry of a two by two contingency table. *J. Amer. Statist. Assoc.* **65** 694–701.

GOOD, I. J. (1976). On the application of symmetric Dirichlet distributions and their mixtures to contingency tables. *Ann. Statist.* **4** 1159–1189.

GOODMAN, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **74** 537–552.

DEPARTMENT OF STATISTICS
CARNEGIE-MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213