LEO A. GOODMAN[1]

University of Chicago

NEW METHODS FOR THE ANALYSIS OF TWO-WAY CONTINGENCY TABLES: AN ALTERNATIVE TO DIACONIS AND EFRON

I would like to congratulate Persi Diaconis and Bradley Efron on their very interesting and stimulating paper (1985). Their paper introduces into the literature on the analysis of contingency tables new ideas and methods, and a new level of mathematical sophistication and depth, pertaining to the interpretation of the chi-square statistic. I appreciate very much the merits of their approach; and I would like here to suggest an alternative approach to the analysis of contingency tables, and to indicate some of the merits of this alternative. I shall illustrate this alternative approach by applying it in an analysis of the same two contingency tables (Tables 1 and 2) that were used by Diaconis and Efron to illustrate their methods.

I begin with Table 1, a 4×4 cross-classification of 592 subjects with respect to eye color (the row classification) and hair color (the column classification). For this table, Diaconis and Efron obtain a chi-square value of 138.29, with $3 \times 3 = 9$ degrees of freedom, using the usual goodness-of-fit (i.e., Pearsonian) chi-square statistic for testing the usual model H_0 of independence between the row classification and column classification. They then proceed to provide new and interesting interpretations of this statistic. The alternative approach that I wish to suggest here, for analyzing contingency tables of this kind, leads me, when analyzing this 4×4 table, to modify the usual model H_0 of independence by introducing 2 parameters pertaining to the association in the table, and the model thus obtained (called here model H') turns out to reduce the goodness-of-fit chi-square value from 138.29 to 10.48, with $9 - 2 = 7$ degrees of freedom (d.f.). This reduction of 92 percent in the chi-square value, with the introduction of only 2 association parameters in model H', is quite dramatic. The 2 parameters provide a parsimonious description of the association in this table; the estimate of these association parameters describes the magnitude of this association; the relatively small chi-square value indicates that the data in this table are congruent (more or less) with H'; the 7 d.f. corresponding to H' indicate that this model makes 7 distinct assertions about the association in the table; and the relatively small chi-square value indicates that the data are congruent (more or less) with those 7 assertions.[2]

[1] Support for this work was provided in part by National Science Foundation Grant SES-8303838. This work was carried out while the author was a Fellow of the Center for Advanced Study in the Behavioral Sciences, with financial support provided in part by National Science Foundation Grant BNS-8011494 to the Center.

[2] Model H' is an example of the kind of model introduced in my work on "association models" (see, e.g., 1979, 1981), as are the other models that will be brought forward in this discussion (viz., models H'', M', M'', and the other association models in Tables A and B). Due to space constraints imposed by the Editor on this discussion, the details about these particular models, and their application in the present context, will be discussed more fully elsewhere.

TABLE 1

Cross-classification of 592 subjects according to their eye color and hair color

	Hair Color				
Eye Color	**Black**	**Brunette**	**Red**	**Blond**	**Total**
Brown	68	119	26	7	220
Blue	20	84	17	94	215
Hazel	15	54	14	10	93
Green	5	29	14	16	64
Total	108	286	71	127	592

TABLE 2

Cross-classification of 25,263 Swedish families according to number of children per family and yearly income

Number of Children	Yearly Income Units of 1000 Kronor				
	0–1	**1–2**	**2–3**	**3+**	**Total**
0	2161	3577	2184	1636	9558
1	2755	5081	2222	1052	11110
2	936	1753	640	306	3635
3	225	419	96	38	778
≥4	39	98	31	14	182
Total	6116	10928	5173	3046	25263

If I allowed myself to be carried away by the general approach that I wish to suggest here, when analyzing this 4×4 table, I might further modify H′ by introducing some additional parameters pertaining to the association in the table, and the model thus obtained (called here model H″) turns out to reduce the goodness-of-fit chi-square value from 138.29 to 10.48 to 0.82, with $9 - 6 = 3$ d.f. Model H″ would make only 3 (rather than 7) distinct assertions about the association in the table, and the tiny chi-square value would indicate that the data are very congruent with the 3 assertions. After the introduction in the preceding paragraph of model H′, which the data fit rather well (with 2 association parameters), the introduction then of model H″, which the data fit very well (with 4 additional association parameters), might be regarded as gilding the lily. (Perhaps I should call H″ the Gilded-Lily model.) I will say more about H′, H″, and other related models below.

Before proceeding with this development, let me comment, however, on the statement by Diaconis and Efron that "with small sample sizes we can often find models to give a satisfactory fit. With large samples, no model fits." (When making this statement, they refer to Joseph Berkson's classic paper (1938) on the chi-square test.) Table 1, which was analyzed in their paper, and considered very briefly in the preceding two paragraphs, described a sample of 592 subjects, and Table 2 in their paper described a sample of 25,263 subjects. Table 1, with $n = 592$, would not usually be viewed as a small sample in the present context;

and Table 2, with $n = 25{,}263$, would be viewed as a large sample. I shall now comment briefly on the analysis of Table 2.

This table is a 5×4 cross-classification of 25,263 families with respect to number of children per family (the row classification) and yearly income (the column classification). For this table, Diaconis and Efron obtain a goodness-of-fit chi-square value of 568.57, with $4 \times 3 = 12$ d.f., for testing the usual model of independence between the row classification and column classification; and they then proceed to provide new and interesting interpretations of this statistic. (The independence model for Table 2 will be called here model M_0 to distinguish it from the corresponding model H_0 for Table 1.) The alternative approach that I wish to suggest here, for analyzing contingency tables of this kind, leads me, when analyzing this 5×4 table, to modify model M_0, which has 4×3 d.f., to an association model M', which has 3×2 d.f.; and this model turns out to reduce the goodness-of-fit chi-square value from 568.57 to 10.97, with 6 d.f. This reduction of 98 percent in the chi-square value is again quite dramatic. Model M' makes 6 distinct assertions about the association in the table, and the relatively small chi-square value indicates that the data are congruent (more or less) with the 6 assertions.[3]

Model M' can be modified further by introducing 2 additional parameters pertaining to the association in this 5×4 table, and the model thus obtained (called here model M'') turns out to reduce the goodness-of-fit chi-square value from 568.57 to 10.97 to 3.64, with 4 d.f. (and with $n = 25{,}263$!). Model M'' would make only 4 (rather than 6) distinct assertions about the association in the table, and the small chi-square value would indicate that the data are congruent with the 4 assertions. After the introduction in the preceding paragraph of model M', which the data fit rather well, the introduction then of model M'', which the data fit really better (with 2 additional association parameters), might be regarded as icing the cake.

Now see Table A, which summarizes the results described above for contingency tables 1 and 2, using both the goodness-of-fit chi-square and the chi-square based on the likelihood-ratio. (Both these chi-square statistics are considered by Diaconis and Efron.) Since there is a greater transition (with respect to d.f.) from H' to H'' for Table 1 than there is from M' to M'' for Table 2, to ease this transition, I have included additional association models for Table 1 in Table A.

Before proceeding further, let me comment briefly on the statement by Diaconis and Efron that the methods they have suggested "are not intended to replace a careful structural investigation," and on their related statement mentioning specifically "log-linear modeling, correspondence analysis, and other structural models." The general approach that I am suggesting here is based on the use of a class of "association models" introduced in my recent work (see, e.g.,

[3] For those familiar with the notation in my earlier articles (see, e.g., 1981), model M' is the RC association model, and the 6 distinct assertions about the association in the table can be expressed in terms of the odds-ratios in the 2×2 subtables formed from the 5×4 table. Further details about this particular model can be found in my earlier articles; and, due to space constraints, application of the model in the present context will be discussed more fully elsewhere.

TABLE A
Association models applied to Contingency Tables 1 and 2

Association Models	Degrees of Freedom	Goodness-of-Fit Chi-Square	Likelihood-Ratio Chi-Square
For Contingency Table 1			
H_0	9	138.29	146.44
$H_1 = H'$	7	10.48	10.08
H_2	5	7.18	6.36
H_3	4	5.22	4.88
$H_4 \doteq H''$	3	0.82	0.83
For Contingency Table 2			
M_0	12	568.57	569.42
M'	6	10.97	11.05
M''	4	3.64	3.76

1979, 1984, 1985). This class of models is different from the class of log-linear models (although the classes do overlap), and the approach is also different from correspondence analysis. With respect to the relationship to log-linear modeling, three of the association models that I have mentioned in this discussion (viz., H'', M', and M'') are not equivalent to log-linear models, and one of the association models (viz., H') can be reexpressed as a log-linear model. With respect to correspondence analysis, I shall now mention briefly some differences between the association models approach and correspondence analysis.

Correspondence analysis, as it is now practiced, is a descriptive tool, which does not usually apply methods of statistical inference. The need to develop the possible inferential aspects of correspondence analysis has been noted (e.g., Nishisato, 1980; Aitkin, 1982); and I have presented, in the 1983 Rietz Memorial Lecture, statistical methods that should help to meet this need when correspondence analysis is applied to contingency tables (see Goodman, 1985). (For related, but different, methods, see, e.g., Lebart, Morineau, and Warwick, 1984; Greenacre, 1984; Gilula and Haberman, 1984.) The Rietz Lecture also compared these statistical methods for correspondence analysis with the statistical methods appropriate for the analysis of association models. (For further details, see the lecture.) With respect to the present brief discussion, I would point out that the application of correspondence analysis to Tables 1 and 2 does not lead to results that are as satisfactory as those obtained with association models. My 1981 article and the Rietz Lecture described conditions under which application of the correspondence analysis approach (and the equivalent canonical correlation approach) would lead to results that are similar, in some respects, to those obtained with the association models approach. (The four different contingency tables that were analyzed from this perspective in these two earlier papers were used to illustrate these similarities.) Tables 1 and 2, on the other hand, do not satisfy the conditions described in those two papers. To illustrate only some of the differences in the results obtained when Tables 1 and 2 are analyzed using the correspondence analysis approach and the association model approach, I have included Table B, which gives the chi-square values that turn up (with 4

TABLE B

Correspondence analysis model and related association model applied to Contingency Tables 1 and 2

Models	Degrees of Freedom	Goodness-of-Fit Chi-Square	Likelihood-Ratio Chi-Square
For Contingency Table 1			
Correspondence Analysis RC	4	14.90	14.17
Association Model RC	4	8.67	8.08
For Contingency Table 2			
Correspondence Analysis RC	6	18.54	19.02
Association Model RC	6	10.97	11.05

d.f. for Table 1 and 6 d.f. for Table 2) when correspondence analysis is applied to the two contingency tables (with the appropriate statistical methods described in the Rietz Lecture) using only the first principal component, and when a related association model is applied to these tables. The contingency tables do not fit the correspondence analysis models, but they do fit (more or less) the related association models. (Comparing Tables A and B, we also find that, with respect to Table 1, the data fit the particular association model, with 4 d.f., in Table A (viz., H_3) better than they do the related association model, with 4 d.f., in Table B. And, with respect to Table 2, the particular association model, with 6 d.f., in Table A (viz., M′) is the same as the related association model in Table B.)[4]

I have not yet described in this discussion what the association models tell us about the association in the two contingency tables. These models tell us a lot; but space constraints on this brief discussion mean that these results will have to be reported elsewhere. Nevertheless, using the perspective introduced by the association models, with respect to Table 1, let me note here only that the data can be used to determine (in a clearly defined sense) an appropriate ordering of the row categories (eye color) and an appropriate ordering of the column categories (hair color); and, with the table as displayed by Diaconis and Efron (and also as displayed by Snee, 1974), the ordering is correct for the column classification, but incorrect for the row classification. With respect to Table 2, let me note here only that the data can be used to determine that the relationship between the row classification (number of children per family) and the column classification (yearly income) is not monotonic, which contradicts (to some extent) the statement that there is "an obvious systematic trend," which was made in an earlier discussion of Table 2 by Abildgård (1974). And, with respect to the comparison of Tables 1 and 2, let me note here only that the association model (and the correspondence analysis model) applied in Table B can be used to show that the data in Table 2 are closer, in several respects, to the independence model than are the data in Table 1. (The results pertaining to the comparison of Tables 1 and 2, which are obtained with the perspective suggested here, provide additional insights into the corresponding result obtained by Diaconis and Efron.)

[4] For those familiar with the notation in my earlier articles, with respect to Table 1, the association models with 4 d.f. in Tables A and B are R + C and RC models, respectively. And, with respect to Table 2, the association model with 6 d.f. is the RC model in both Tables A and B.

With respect to the application of the association models, to further increase confidence in their utility, I have begun to study the chi-square values reported in Table A, using tools from the cross-validation, jackknife, and bootstrap literature. The models in Table A were selected to fit the data (but were limited to the class of models described in, e.g., my 1981 article), and so an evaluation of the chi-square values thus obtained will be of interest. This too I shall report on elsewhere.

In closing, let me note that there are similarities between the association models considered here for the analysis of a two-way contingency table and nonadditivity models in the analysis of a two-way array with continuous data, of the kind considered in, e.g., Mandel (1971), Johnson and Graybill (1972), and Tukey (1949, 1977). The development of methods appropriate in the contingency table context can lead to the development of some related methods appropriate for the analysis of nonadditivity models in a two-way array with continuous data, and these related methods will make a new contribution to the statistical tools now available for the analysis of nonadditivity with the analysis of variance; see Goodman and Haberman (1985).

One final comment: In closing their paper, Diaconis and Efron note that Harold Hotelling's volume test (in a regression context), which was developed in his 1939 article, is closely related to their volume test (in the contingency table context); and they mention that the "statistical content (of Hotelling's article) seems to have gone largely ignored (except for scattered applications, as in Efron, 1971)." I expect that much more attention will be paid to the statistical content in their paper than was paid to the corresponding content in Hotelling's article. And with respect to the alternative suggested in the present discussion (viz., the association models approach developed in my recent work): judging from the reception of my earlier work on log-linear models (e.g., 1969, 1970, 1971, 1978), I suspect that the statistical content of the association models approach will not go ignored. The association models provide meaningful and parsimonious descriptions of the association in many different contingency tables, and data obtained in many different contexts are found to be congruent with these models (see Goodman, 1984, 1985). These models, it would appear, are chock full of statistical content.

REFERENCES

ABILDGÅRD, F. (1974). Discussion of "The notion of redundancy and its use as a quantitative measure of the discrepancy between a statistical hypothesis and a set of observational data," by Per Martin-Löf. *Scand. J. Statist.* **1** 13–15.

AITKIN, M. (1982). Review of *Analysis of Categorical Data; Dual Scaling and Its Applications*, by S. Nishisato. *J. Roy. Statist. Soc. Ser. A* **145** 513–516.

BERKSON, J. (1938). Some difficulties of interpretation encountered in the applcation of the chi-square test. *J. Amer. Statist. Assoc.* **33** 526–542.

CLOGG, C. C. (1982a). Using association models in sociological research: Some examples. *Amer. J. Sociology* **88** 114–134.

CLOGG, C. C. (1982b). Some models for the analysis of association in multiway cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **77** 803–815.

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.

EFRON, B. (1971). Does an observed sequence of numbers follow a simple rule? (Another look at Bode's law). *J. Amer. Statist. Assoc.* **66** 552–559.
GILULA, Z. and HABERMAN, S. (1984). Canonical analysis of contingency tables by maximum likelihood. Unpublished manuscript.
GOODMAN, L. A. (1969). On partitioning χ^2 and detecting partial association in three-way contingency tables. *J. Roy. Statist. Soc. Ser. B* **31** 486–498.
GOODMAN, L. A. (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *J. Amer. Statist. Assoc.* **65** 226–256.
GOODMAN, L. A. (1971). The analysis of multidimensional contingency tables: Stepwise procedures and direct estimation methods for building models for multiple classifications. *Technometrics* **13** 33–61.
GOODMAN, L. A. (1978). *Analyzing Qualitative/Categorical Data: Log-Linear Models and Latent-Structure Analysis.* Abt Books, Cambridge, Mass.
GOODMAN, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **74** 537–552.
GOODMAN, L. A. (1981). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **76** 320–334.
GOODMAN, L. A. (1984). *The Analysis of Cross-Classified Data Having Ordered Categories.* Harvard University Press, Cambridge, Mass.
GOODMAN, L. A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Ann. Statist.* **13** 10–69.
GOODMAN, L. A. and HABERMAN, S. J. (1985). On the analysis of non-additivity in two-way analysis of variance. Unpublished manuscript.
GREENACRE, M. J. (1984). *Correspondence Analysis.* Academic, New York.
HOTELLING, H. (1939). Tubes and spheres in n-spaces, and a class of statistical problems. *Amer. J. Math.* **61** 440–460.
JOHNSON, D. E. and GRAYBILL, F. A. (1972). An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc.* **67** 862–868.
LEBART, L., MORINEAU, A. and WARWICK, K. M. (1984). *Multivariate Descriptive Statistical Analysis.* Wiley, New York.
MANDEL, J. (1971). A new analysis of variance model for non-additive data. *Technometrics* **13** 1–18.
NISHISATO, S. (1980). *Analysis of Categorical Data: Dual Scaling and Its Applications.* University of Toronto Press, Toronto.
SNEE, R. (1974). Graphical display of two-way contingency tables. *Amer. Statist.* **28** 9–12.
TUKEY, J. W. (1949). One degree of freedom for nonadditivity. *Biometrics* **5,** 232–224.
TUKEY, J. W. (1977). *Exploratory Data Analysis.* Addison-Wesley, Reading, Mass.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS 60637

TOM LEONARD[1]

University of Wisconsin-Madison

This exciting and imaginative paper promises substantial impact upon the practical application of contingency table methodology. It highlights the difficulty for the user in making an overall judgement concerning a reduced model, by reference to the tricky (and dare I say virtually impossible?) interpretation,

[1] This work was supported under Contract N00014-83-C-0514 between the Office of Naval Research and the University of Iowa, Melvin R. Novick, Principal Investigator, while the author was visiting the Lindquist Center for Measurement, University of Iowa.