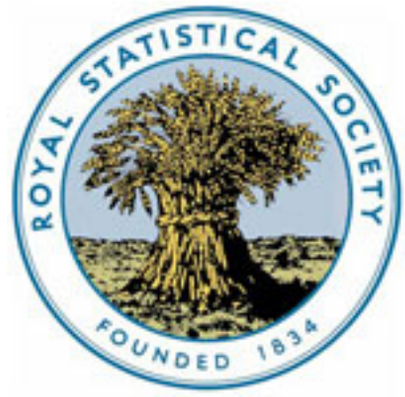

Appl. Statist. (1974),
23, *No.* 3. p. 405

Identification of the Sources of Significance in Two-way Contingency Tables

By MORTON B. BROWN

University of California, Los Angeles†

[Received April 1973. Final revision February 1974]

SUMMARY

When the χ^2 test for independence between rows and columns in a two-way contingency table is significant, a procedure is described for sequentially identifying those cells which contribute heavily to the χ^2 statistic. A 14×14 contingency table is used to compare the graphical technique of Fienberg (1969) with this method. An algorithm is presented to estimate the expected values of each cell under the model of quasi-independence.

Keywords: TWO-WAY CONTINGENCY TABLES; OUTLIERS; MISSING VALUES; QUASI-INDEPENDENCE

1. INTRODUCTION

GOODMAN (1968, 1969, 1971) discusses testing the hypothesis of quasi-independence in a subset of a two-way contingency table. A quasi-independent subset is a subset for which the cell frequencies a_{ij} can be fitted by the model $E(a_{ij}) = \alpha_i \beta_j$ where α_i is a function only of the rows and β_j only of the columns. When some cells are empty *a priori* or missing, the hypothesis of quasi-independence between the rows and columns for the remaining cell frequencies is tested by a χ^2 statistic with an adjustment to its degrees of freedom.

When the null hypothesis of independence is rejected in a table with possibly empty cells, the investigator may want to identify those cells which contribute most to the χ^2. Fienberg (1969) proposes a graphical technique based on half-normal plots to identify those 2×2 sub-tables for which the null hypothesis of independence should be rejected. Two apparent drawbacks to his approach are its dependence on the original order of the rows and columns and the identification of too many cells.

Haberman (1973) suggests the use of adjusted standardized deviations to measure the disparity between an observed cell frequency and its expectation. By "adjusted" is meant the ratio of the difference between the observed and expected frequencies to the standard deviation of the difference. He then plots the adjusted deviations on a full normal plot to identify cells that appear deviant. The expected values are estimated using all the cells. Therefore, extreme cells can bias the expected values sufficiently to make it difficult to draw inferences about other than the most extreme cells.

A stepwise algorithm is described which identifies and eliminates cells inconsistent with a quasi-independent model estimated from the remaining cells. Using a 14×14 contingency table previously analysed by Fienberg (1969), the cells selected for elimination by this approach are compared with those chosen by Fienberg's procedure.

† Author's permanent address: Department of Statistics, Tel-Aviv University, Tel-Aviv, Israel.

Both Goodman (1968) and Fienberg (1969) give methods for fitting quasi-independent models. That of Fienberg iterates over all observed cells, while Goodman's cycles over row and column totals. An algorithm is presented that iterates over the empty (or eliminated) cells. When a small number of cells are empty, this procedure requires fewer computations than either of the other two.

2. Elimination of a Single Cell

Let $\{a_{ij}, i = 1, \ldots, R;\ j = 1, \ldots, C\}$ be the observed frequency counts in the $R \times C$ cells in a two-way contingency table. Let

$$r_i = \sum_j a_{ij}, \quad c_j = \sum_i a_{ij}, \quad \text{and} \quad N = \sum_i \sum_j a_{ij}$$

be the row marginals, column marginals and the total frequency of the table. The sampling is assumed to be multinomial conditional on the total frequency N or on either the row or the column marginals. Under the assumption of independence between the rows and the columns, the estimated expected value of cell (i, j) is

$$E_{ij} = r_i c_j / N. \tag{1}$$

The χ^2 test for independence between the rows and the columns is

$$\chi^2 = \sum_i \sum_j d_{ij}^2 \quad \text{where} \quad d_{ij} = (a_{ij} - E_{ij})/\sqrt{E_{ij}}. \tag{2}$$

Goodman (1968) gives maximum-likelihood estimates (M.L.E.) of the expected values for all observed cells when one cell, the (I,J)th, is assumed to be empty *a priori*. We give a derivation of these formulae which is generalized to more than one missing cell in Section 5. Let the observed value a_{IJ} of cell (I,J) be replaced by its M.L.E. from the model of quasi-independence for a table in which only cell (I,J) is assumed empty *a priori*. Then find the expected values of all cells by (1) as if the table had no empty cells. The expected value of cell (I,J) must be equal to E^*_{IJ} (Watson, 1956); that is,

$$E^*_{IJ} = (r_I + E^*_{IJ} - a_{IJ})(c_J + E^*_{IJ} - a_{IJ})/(N + E^*_{IJ} - a_{IJ}).$$

Solving for E^*_{IJ} yields

$$E^*_{IJ} = (r_I - a_{IJ})(c_J - a_{IJ})/(N - r_I - c_J + a_{IJ}). \tag{3}$$

After substituting for E^*_{IJ}, the estimated expected values of the other cells are

$$\left.\begin{aligned} E^*_{Ij} &= (r_I - a_{IJ})\, c_j/(N - c_J), \quad j \neq J, \\ E^*_{iJ} &= r_i(c_J - a_{IJ})/(N - r_I), \quad i \neq I, \\ \text{and} & \\ E^*_{ij} &= r_i c_j(N - r_I - c_J + a_{IJ})/\{(N - r_I)(N - c_J)\}, \quad i \neq I, j \neq J. \end{aligned}\right\} \tag{4}$$

Therefore, in a table with only one empty cell, the expected values of the observed cells under a quasi-independent model can be estimated by replacing the value in that cell by E^*_{IJ} (3) and proceeding as if all cells are present.

Assuming cell (I,J) is empty, the χ^2 statistic χ^2_{IJ} can be calculated using formulae (3) and (4) for the expected values. Then

$$\chi^2_{IJ} = \frac{(N-r_I)(N-c_J)}{(N-r_I-c_J+a_{IJ})} S_{++} - N + \left(\frac{Na_{IJ}-r_I c_J}{N-r_I-c_J+a_{IJ}}\right)\left(\frac{(N-c_J)}{r_I(r_I-a_{IJ})} S_{I+}\right.$$

$$\left. + \frac{(N-r_I)}{c_J(c_J-a_{IJ})} S_{+J} - \frac{a_{IJ}\{r_I c_J(N-r_I-c_J+a_{IJ}) - a_{IJ}(a_{IJ}N - r_I c_J)\}}{r_I(r_I-a_{IJ})\, c_J(c_J-a_{IJ})}\right), \quad (5)$$

where

$$S_{++} = \sum_i \sum_j a_{ij}^2/r_i c_j,$$

$$S_{I+} = \sum_j a_{Ij}^2/c_j$$

and

$$S_{+J} = \sum_i a_{iJ}^2/r_i.$$

That cell which, when assumed empty *a priori*, produces the minimum χ^2_{IJ} is the most divergent in terms of the test for independence.

The difference between χ^2_{IJ} and the χ^2 obtained for the original table can be expressed as

$$\delta_{IJ} = \Delta\left\{S_{++} - 1 - \frac{(N+\Delta)}{c_J(c_J+\Delta)} S_{+J} - \frac{(N+\Delta)}{r_I(r_I+\Delta)} S_{I+}\right.$$

$$\left. + \frac{(N+\Delta)}{(r_I+\Delta)(c_J+\Delta)}\left(\frac{\Delta a_{IJ}^2}{r_I c_J} + 2a_{IJ} + \Delta\right)\right\}, \quad (6)$$

where

$$\Delta = \text{change in value for cell } (I,J)$$

$$= E^*_{IJ} - a_{IJ}$$

$$= (r_I c_J - a_{IJ} N)/(N - r_I - c_J + a_{IJ})$$

after substituting for E^*_{IJ} (3). The above expression for δ_{IJ} can be used to evaluate the net change in the χ^2 by a change (Δ) in any cell value.

3. A Stepwise Searching Procedure for Identifying a Quasi-independent Subset

When the hypothesis of independence in a two-way contingency table is rejected, the identification of those cells that contribute the most to the significant χ^2 can be of value in drawing inferences about the table. The following stepwise procedure selects sequentially the cell whose elimination causes the greatest reduction in the value of the χ^2 statistic. A cell which is so selected is thereafter treated as empty *a priori*.

For each cell in the original table calculate χ^2_{IJ} (5). Select the cell whose elimination produces the smallest χ^2. Replace the observed value of that cell by E^*_{IJ} (3). In subsequent steps, calculate χ^2_{IJ} for all non-selected cells by assuming that the observed values for the selected cells are equal to their expected values under the quasi-independent model. After the cell associated with the minimal χ^2_{IJ} is deleted, refit the expected values for all the cells by a quasi-independent model based on the non-deleted cells (Section 5) and recalculate the χ^2 test for quasi-independence.

Since χ^2_{IJ} for the cell to be deleted is calculated by assuming that previously deleted cells have observed value equal to their expected values at this step, χ^2_{IJ} is not equal to the χ^2 test for quasi-independence obtained after refitting the expected values. This occurs because the expected values of all the cells alter each time an additional cell is eliminated. However, χ^2_{IJ} is a good approximation to the answer. In the example considered in the following section, the two χ^2 tests at any step did not differ by more than 0·3.

As in stepwise regression, there is no guarantee that the first k cells eliminated yield the minimal χ^2 for any such set of k cells. As long as cells are extreme, the order of selection can vary from optimal but all extreme cells are likely to be deleted. After deleting a subset of cells, it is possible to calculate the effect on the χ^2 by reversing the procedure and reincluding any eliminated cell by using equation (6) in which Δ is set equal to the difference between the original observed frequency of that cell and its expected value under the quasi-independent model.

The following stopping rule suggests itself: choose a level of significance α and stop eliminating cells as soon as the significance of the χ^2 statistic exceeds α. If, for the the original frequency table, the hypothesis of independence (or quasi-independence) is appropriate, then with probability $(1-\alpha)$ no cell will be eliminated. This procedure for eliminating cells is similar to the Newman–Keuls multiple-range test (Winer, 1971) in which at any step the cells are ordered by their χ^2_{IJ} rather than by group means. It is continued to a following step only when rejection of the null hypothesis occurs.

4. An Example

Fienberg (1969) reanalyses a 14×14 contingency table containing father/son occupations which appears in Pearson (1904). This historical set of data is presented in Table 1. Since many of the cells are zero, the table is analysed twice: the first time the original cell entries are used; the second time $\frac{1}{2}$ is added to each cell before analysing the table. The results of the stepwise algorithm, applied to both the original and the augmented tables, are shown in Table 2.

For both tables cells are eliminated in similar orders. The χ^2 tests for quasi-independence yield smaller χ^2's when the augmented table is used. This is a reflection of the lack of robustness of the χ^2 statistic to small expected cell frequencies. Therefore, conclusions are drawn from the analysis of the modified data.

The first eight cells selected are all diagonal cells. Three of the next four are off-diagonal. They represent the father–son occupations: crafts–arts; divinity–medicine; and landownership–art. The χ^2 test is no longer significant at the twelfth step, indicating that we should stop no later than that step. The conclusions to be drawn from the above analysis are: sons tend to remain in the same occupational category as their fathers; and, except for very few off-diagonal pairings, when a son switches to another category, his choice appears independent of his father's occupation. The first part of the conclusions needs to be tempered by the change in the marginal totals of the table. In this set of data, which reflects England at the turn of the century, there is a flight by sons from crafts, agriculture, landownership and commerce into the arts, literature, scholarship and science. However, when the son differs in occupation from his father, his choice is in general independent of his father's occupation.

Three possible criteria other than the above (a) for the choice of cells in a stepwise fashion are included in Table 2. They are: (b) delete the cell which has maximum $|d_{IJ}| = |a_{IJ}-E_{IJ}|/\sqrt{E_{IJ}}$; (c) delete the cell which has the largest absolute adjusted

deviation (Haberman, 1973)

$$|d'_{IJ}| = |a_{IJ} - E_{IJ}| \Big/ \left\{ N \frac{r_I}{N}\left(1 - \frac{r_I}{N}\right) \frac{c_J}{N}\left(1 - \frac{c_J}{N}\right) \right\}^{\frac{1}{2}} \tag{7}$$

and (d) delete the cell which has the maximum $|d^*_{IJ}| = |a_{IJ} - E^*_{IJ}|/\sqrt{E^*_{IJ}}$ where E^*_{IJ} is defined in (3). These three are obvious measures of the distance of an observation from its expected value.

TABLE 1

Father/son occupations (*Pearson*, 1904)

Father's occupation	*Son's occupation*														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	*Total*
(1)	28	0	4	0	0	0	1	3	3	0	3	1	5	2	50
(2)	2	51	1	1	2	0	0	1	2	0	0	0	1	1	62
(3)	6	5	7	0	9	1	3	6	4	2	1	1	2	7	54
(4)	0	12	0	6	5	0	0	1	7	1	2	0	0	10	44
(5)	5	5	2	1	54	0	0	6	9	4	12	3	1	13	115
(6)	0	2	3	0	3	0	0	1	4	1	4	2	1	5	26
(7)	17	1	4	0	14	0	6	11	4	1	3	3	17	7	88
(8)	3	5	6	0	6	0	2	18	13	1	1	1	8	5	69
(9)	0	1	1	0	4	0	0	1	4	0	2	1	1	4	19
(10)	12	16	4	1	15	0	0	5	13	11	6	1	7	15	106
(11)	0	4	2	0	1	0	0	0	3	0	20	0	5	6	41
(12)	1	3	1	0	0	0	1	0	1	1	1	6	2	1	18
(13)	5	0	2	0	3	0	1	8	1	2	2	3	23	1	51
(14)	5	3	0	2	6	0	1	3	1	0	0	1	1	9	32
Total	84	108	37	11	122	1	15	64	69	24	57	23	74	86	775

The categories are: (1) army, (2) art, (3) teacher, clerk, civil servant, (4) crafts, (5) divinity, (6) agriculture, (7) landownership, (8) law, (9) literature, (10) commerce, (11) medicine, (12) navy, (13) politics and court, (14) scholarship and science.

The orders of the selection of cells by the criteria vary although all choose the same first 10 cells. The last criterion $|d^*_{IJ}|$ has a similar order of selection as χ^2_{IJ} for the 10 cells. The order of selection by Haberman's criterion is closer to that of χ^2_{IJ} than that of $|d_{IJ}|$. After step 10, the three criteria differ in choice from χ^2_{IJ}. It appears that there are alternate paths of deletion which reduce the χ^2 in approximately similar amounts. The major reductions in the χ^2 obtained by deleting cells occurred by step 10. At this point, there are several different quasi-independent subsets that can be identified which differ only in one or two cell elements.

The 44 cells eliminated by Fienberg are shown in Table 3. For each cell the expected value and the standardized differences are estimated from a quasi-independent model fitted to all but the 44 cells. Assuming small cell probabilities and the hypothesis of independence, each d^*_{ij} is approximately standard normal. Probabilities of rejecting this null hypothesis are also given in Table 3. Ten of the 13 cells that are significant at the 0·001 level are included in the first 12 cells eliminated by our algorithm. More than half (28 of the 44 cells) do not differ significantly from their expected values as estimated by fitting the model of quasi-independence.

TABLE 2

Stepwise elimination of cells in Table 1 and the resulting χ^2 test for quasi-independence

	Original data			*Data augmented by* $\frac{1}{2}$								
Step	*Cell eliminated* (*a*)†	χ^2	*Probability*	*Cell eliminated* (*a*)†	χ^2	*Probability*	*Cell eliminated* (*b*)†	χ^2	*Cell eliminated* (*c*)†	χ^2	*Cell eliminated* (*d*)†	χ^2
0		1005·4	0·000		877·5	0·000		877·5		877·5		877·5
1	2 2	721·5	0·000	2 2	614·8	0·000	2 2	614·9	2 2	614·9	2 2	614·9
2	5 5	608·1	0·000	5 5	507·7	0·000	1 1	518·7	1 1	518·7	5 5	507·7
3	1 1	510·0	0·000	1 1	418·9	0·000	11 11	435·1	5 5	418·9	1 1	418·9
4	11 11	426·3	0·000	11 11	345·0	0·000	5 5	345·0	11 11	345·0	11 11	345·0
5	13 13	371·4	0·000	13 13	294·5	0·000	13 13	294·5	13 13	294·5	13 13	294·5
6	12 12	329·3	0·000	12 12	265·4	0·000	12 12	265·4	12 12	265·4	12 12	265·4
7	4 4	304·9	0·000	4 4	247·4	0·000	4 4	247·4	4 4	247·4	4 4	247·4
8	8 8	286·0	0·000	8 8	229·8	0·000	4 2	232·0	4 2	232·0	8 8	229·8
9	4 2	270·0	0·000	4 2	215·0	0·002	8 8	215·0	8 8	215·0	4 2	215·0
10	5 11	256·3	0·000	5 11	202·7	0·011	5 11	202·7	5 11	202·7	5 11	202·7
11	10 10	245·0	0·000	10 10	193·6	0·028	7 13	195·2	7 13	195·2	7 13	195·2
12	7 2	235·4	0·000	7 2	184·7	0·065	7 1	184·8	7 1	184·8	7 1	184·8
13	13 8	226·2	0·000	13 8	176·6	0·123	13 8	177·7	13 8	177·7	10 10	176·4
14	7 9	218·7	0·001	7 9	169·3	0·204	7 7	172·0	10 10	169·6	13 8	169·6
15	7 14	211·8	0·001	7 14	162·3	0·308	7 8	164·0	7 7	164·0	10 2	162·6
16	13 12	205·7	0·003									
17	7 4	200·0	0·005									
18	10 2	194·5	0·010									
19	5 13	189·0	0·017									
20	1 5	183·9	0·027									
21	14 4	180·3	0·037									
22	2 4	171·4	0·082									
23	4 1	166·4	0·119									

† The cell eliminated is the one which:
(*a*) minimizes χ^2_{IJ} (5);
(*b*) maximizes $\left|\begin{matrix} d_{IJ} \\ d'_{IJ} \end{matrix}\right| = |a_{IJ}-E_{IJ}|/\sqrt{E_{IJ}}$;
(*c*) maximizes $\left|\begin{matrix} d_{IJ} \\ d'_{IJ} \end{matrix}\right|$, Haberman's adjusted deviation (7);
(*d*) maximizes $\left|d^*_{IJ}\right| = |a_{IJ}-E^*_{IJ}|/\sqrt{E^*_{IJ}}$.

TABLE 3

Cells eliminated by Fienberg (1969), *their observed and expected values as estimated from a quasi-independent model with all* 44 *cells excluded*

Cell	Observed value a_{ij}	Expected value E^*_{ij}	d^*_{ij}	Cell	Observed value a_{ij}	Expected value E^*_{ij}	d^*_{ij}	Cell	Observed value a_{ij}	Expected value E^*_{ij}	d^*_{ij}
1 1	28	2·49	16·18***	5 5	54	9·94	13·98***	8 8	18	4·88	5·94***
1 2	0	3·01	−1·74	6 8	1	2·18	−0·80	10 10	11	2·99	4·64***
2 1	2	0·98	1·03	6 9	4	3·43	0·31	10 11	6	7·07	−0·40
2 2	51	1·19	45·64***	6 12	2	0·59	1·84	11 10	0	0·72	−0·85
2 3	1	0·68	0·38	6 13	1	1·99	−0·70	11 11	20	1·71	13·98***
3 1	6	4·48	0·72	7 1	17	5·57	4·84***	11 12	0	0·54	−0·73
3 2	5	5·43	−0·19	7 2	1	6·76	−2·21*	12 11	1	0·77	0·26
3 3	7	3·13	2·18*	7 7	6	0·99	5·02***	12 12	6	0·24	11·65***
3 4	0	0·50	−0·70	7 8	11	5·14	2·59**	12 13	2	0·82	1·30
4 1	0	3·24	−1·80	7 9	4	8·06	−1·43	13 12	3	0·72	2·68**
4 2	12	3·93	4·07***	7 12	3	1·39	1·37	13 13	23	2·44	13·17***
4 3	0	2·27	−1·51	7 13	17	4·68	5·69***	13 14	1	4·94	−1·77
4 4	6	0·36	9·40***	8 1	3	5·29	−1·00	14 13	1	2·16	−0·79
4 5	5	5·03	−0·02	8 2	5	6·41	−0·56	14 14	9	4·40	2·19*
5 4	1	0·71	0·35	8 7	2	0·94	1·09				

*** Two-tailed probability < 0·001.
** Two-tailed probability < 0·01.
* Two-tailed probability < 0·05.

5. An Algorithm for Fitting a Sub-table by a Quasi-independent Model

When more than one cell is excluded from a table, the expected values of the remaining calls can be found by solving sets of simultaneous equations. For several patterns of missing cells, Goodman (1968) gives explicit solutions. In general, the following algorithm yields maximum-likelihood estimates of the expected values by iterating such that each cycle modifies the empty or deleted cells.

At the kth iteration, let

$$E_{ij}^{*(k)} = \frac{(r_i^{(k-1)} - E_{ij}^{*(k-1)})(c_j^{(k-1)} - E_{ij}^{*(k-1)})}{N^{(k-1)} - r_i^{(k-1)} - c_j^{(k-1)} + E_{ij}^{*(k-1)}},$$

$$r_i^{(k)} = r_i^{(k-1)} + E_{ij}^{*(k)} - E_{ij}^{*(k-1)},$$

$$c_j^{(k)} = c_j^{(k-1)} + E_{ij}^{*(k)} - E_{ij}^{*(k-1)},$$

$$N^{(k)} = N^{(k-1)} + E_{ij}^{*(k)} - E_{ij}^{*(k-1)}$$

for all the empty cells. To check whether convergence has occurred, test if

$$\underset{\text{empty cells}}{\text{Max}} \left| E_{ij}^{*(k)} - \frac{r_i^{(k)} c_j^{(k)}}{N^{(k)}} \right| < \varepsilon$$

for some suitably chosen ε. We have used $\varepsilon = 0{\cdot}05$. As starting values we use: the original frequency in the cell; 0 if the cell is empty *a priori*; or the expected value calculated by a previous step in a stepwise routine for deletion of cells.

When only one cell is empty, convergence is immediate since the fitted value satisfies equation (3) and, therefore, the criterion for convergence. Otherwise two to five cycles appear to be sufficient to obtain convergence. Our only experience of lack of convergence occurred when the empty or deleted cells formed a pattern such that the expected values of the empty cells were not uniquely determined. This happens when the frequency table is separable; that is, the table can be partitioned into two sub-tables that do not have any rows or any columns in common and which contain all the non-empty or non-eliminated cells.

The row totals, column totals and table total are not held constant but are continually modified during each cycle. However, the partial row, column and table totals summed over the non-empty cells are held constant.

After convergence, replace the observed values of the empty cells by their estimated expected values. The expected values of the remaining cells are then estimated by assuming the table has no empty cells (1). These are the maximum-likelihood estimates (Goodman, 1968; Bishop and Fienberg, 1969).

When the table is not separable, the degrees of freedom associated with the χ^2 statistic can be calculated as follows: let R' and C' be the number of rows and columns respectively that have at least one observed cell and let M be the number of eliminated or missing cells in the $R' \times C'$ sub-table. Then the number of degrees of freedom for the χ^2 goodness-of-fit statistic is $(R'-1)(C'-1)-M$.

6. Concluding Remarks

All four criteria used in Table 2 to select cells sequentially are intuitively plausible. When all frequencies in the table except one (an outlier) can be fitted exactly by a quasi-independent model, only the criteria of minimizing χ_{IJ}^2 or maximizing $|d'_{IJ}|$ always properly identify the cell containing the outlier. The former criterion is zero

only when the quasi-independent model fits exactly all but the (I,J)th frequency. Therefore the minimum occurs for the cell with the outlier. Haberman's adjusted deviation $|d'_{IJ}|$ can be shown to be a maximum for the deviant cell. When the expected value of the cell containing the outlier is large relative to the expected value of another cell in the table, the criterion $|d_{IJ}|$ can fail to identify the outlier. When the outlier differs by a small amount from its expected value according to the quasi-independent model and its frequency is less than that in another cell, $|d^*_{IJ}|$ can select the wrong cell.

The correct identification of a single cell, when it is the only cell that does not fit a quasi-independent model, is a far simpler problem than identifying a subset of cells, all of which deviate from a model fitted to the other cells. An optimal solution to this latter problem is yet to be found. As in stepwise regression, the sequential procedure described here does not necessarily identify that subset whose exclusion will reduce the χ^2 by the maximum amount. In examples where some of the selected cells have a common property, such as lying on the diagonal or in the same row, we have found it helpful to first eliminate all the cells with the same property and then do the analysis again.

Acknowledgements

The author is grateful to R. I. Jennrich, J. W. Frane and the referees for their helpful comments. This research was supported by NIH Special Resources Research Grant RR-3.

References

Bishop, Y. M. M. and Fienberg, S. E. (1969). Incomplete two-dimensional contingency tables. *Biometrics*, **25**, 119–128.

Fienberg, S. E. (1969). Preliminary graphical analysis and quasi-independence for two-way contingency tables. *Appl. Statist.*, **18**, 153–168.

Goodman, L. A. (1968). The analysis of cross-classified data: independence, quasi-independence, and interactions in contingency tables with or without missing entries. *J. Amer. Statist. Ass.*, **63**, 1091–1131.

—— (1969). How to ransack social mobility tables and other kinds of cross-classification tables. *Amer. J. Sociol.*, **75**, 1–40.

—— (1971). A simple simultaneous test procedure for quasi-independence in contingency tables. *Appl. Statist.*, **20**, 165–177.

Haberman, S. J. (1973). The analysis of residuals in cross-classified tables. *Biometrics*, 29, 205–220.

Pearson, K. (1904). On the theory of contingency and its relation to association and normal correlation. Reprinted in 1948 in *Karl Pearson's Early Statistical Papers*, pp. 443–475. Cambridge: University Press.

Watson, G. S. (1956). Mixed and "mixed-up" frequencies in contingency tables. *Biometrics*, **12**, 47–50.

Winer, B. J. (1971). *Statistical Principles in Experimental Design*, 2nd ed. New York: McGraw-Hill.