Maximum Likelihood Estimation and Model Selection for Incomplete Contingency Tables

Fatma A. Abdelaty

Faculty of Commerce
Mansoura University

SUMMARY :

Many contingency tables have some cells that contain structural zeros that must remain empty under any fitted model. This article focuses on the problem of obtaining maximum likelihood estimates (MLE) for the parameters of log-linear models under this type of incomplete tables. The appropriate systems of equations are presented and the generalization of the proportional fitting (PF) algorithm of Deming and Stephen [5] is suggested as one of the possible methods for solving them. The algorithm has certain advantages but the convergence tends to be somewhat slower than for other alternatives. Tests of fit for log-linear models for incomplete tables are considered. The data for patients with strokes (Bishop and Fienberg [1]) are used to illustrate the procedures discussed in the articel.

KEY WORDS: Contingency tables; Maximum likelihood estimation; Nested medels; PF algorithm; Quasi-log-linear model; Separability.

1. INTRODUCTION

When analyzing sample tables of counts, we encounter two types of empty cells: (a) sampling zero cell, is due to sampling variability, and the relative smallness of the cell probability. At least in principle, by increasing the sample size sufficiently we can make sampling zeros disappear, and (b) structural zero cell, is known a priori to have a zero value.

Structural zero cell occur in several different context. Observations for certain cells in a contingency table are often truncated or not reported (Goodman [9], Watson [14]). At other times, certain combinations are impossible, and zero probability is attached to these cells (Kastenbaum [10]). For example, when there is an underlying order for the categories in each of two or more variables, this ordering may constrain certain cells to be zero a priori (Bishop and Fienberg [2], Chen et al. [3], Mantel and Halperin [7]). To illustrate, in an analysis of scores in games, if one variable is the winning score and another the losing score, then the cells with losing scores exceeding winning scores are all zero a priori. We may wish to fit a parametric model to one set of observed cell counts within a table and a second model to the remaining cells (Fienberg [6], Goodman [9],[8], Savage and Deutsch [13]). In such situations, maximum likelihood estimation procedures lead us to treat these sets of expected cell counts as being the nonzero entries in two separate incomplete tables.

The proportional fitting (PF) algorithm is applied here to obtain maximum likelihood estimates (MLE's) for the expected cell frequencies for incomplete tables. Furthermore, for nested sequences of models, we partition the likelihood ratio goodness-of-fit statistic into additive components. Tests of fit for loglinear models are considered. As a result, we suggest fitting the model of Quasi independence.

A brief review of the analysis of complete contingency tables is presented in the next section. Section 3 is devoted to quasi log linear models for incomplete multiway tables. In Section 4, the concept of seperability is introduced. Section 5 is devoted to MLE's for incomplete multiway tables. The data are analyzed in Section 6.

2. COMPLETE CONTINGENCY TABLES

We consider a three-dimensinal contingency table I x J x K, where the indices pertain to the variables A,B,C, respectively. For the complete-table case we let x_{ijk} be the frequency in cell (i,j,k) and $m_{ijk} = E(x_{ijk})$ be its expected value. Further, we let N denote the sample size and $p_{ijk} = m_{ijk}/N$. Then, in the notation of Bishop, Fienberg, and Holland [1], the saturated model is written as

$$\log m_{ijk} = u+u_{1(i)}+u_{2(j)}+u_{3(k)}+u_{12(ij)}+u_{13(ik)}+u_{23(jk)}+u_{123(ijk)}, \qquad (2.1)$$

where the subscripted u terms sum to zero when summed over any subscript. More parsimonious models are postulated by setting appropriate u terms equal to zero.

The sampling schemes that generate the data are assumed to be (a) independent Poisson distributions, (b) multinomial, or (c) product multinomial distributions. It is well known that if the sufficient statistics include the marginals that are fixed by design in the product multinomial distribution, the three sampling schemes lead to the same MLE's (e.g., Haberman [10], Fienberg [7]).

MLE's for the vector of parameters m (or. P) can be obtained by equating the estimated and observed marginals that correspond to the u terms in the model.

For example, for the model of independence

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)},$$

the minimal sufficient statistics are x_{i++} , i=1, ..., I ; x_{+j+} , j=1,...,J ; and x_{++k} , k=1,...,K . (Indices summed over are replaced by +.) The likelihood equations for that model are

$$m_{i++} = x_{i++}, \quad i = 1, \ldots, I$$

$$m_{+j+} = x_{+j+}, \quad j = 1, \ldots, J$$

$$m_{++k} = x_{++k}, \quad k = 1, \ldots, K.$$

3. QUASI-LOG-LINEAR MODELS

We let S be the set of all cells in an incomplete IxJxK three-way array that consists of all cells not containing structural zeros, and m_{ijk} the expected number of individuals in the (i,j,k) cell, where m_{ijk} =0 for (i,j,k) $\notin$ S. We can specify the most general log-linear model for those cells in S, i.e.,

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} \quad (3.1)$$

$$+ u_{23(jk)} + u_{123(ijk)},$$

for (i,j,k) ε S, where the u-terms are deviations and sum to zero over each included variable, e.g.,

$$\Sigma_i \, \delta_i^{(23)} \, u_{1(i)} = \Sigma_i \, \delta_{ij}^{(3)} \, u_{12(ij)} = \Sigma_i \, \delta_{ik}^{(2)} \, u_{13(ik)} \tag{3.2}$$

$$= \Sigma_i \, \delta_{ijk} \, u_{123(ijk)} = 0,$$

with

$$\delta_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \in S \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

$$\delta_{ij}^{(3)} = \begin{cases} 1 & \text{if } \delta_{ijk} = 1 \quad \text{for some } k \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

$$\delta_i^{(23)} \quad \begin{cases} 1 & \text{if } \delta_{ijk} = 1 \quad \text{for some } (j,k) \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

and similar definitions for $\delta_{ik}^{(2)}$, $\delta_{jk}^{(1)}$, $\delta_j^{(13)}$, and $\delta_k^{(12)}$. We note that (3.2) includes some u-terms not found in (3.1), i.e., those preceded by zero values of δ_{ijk} . We set those u-terms in (3.2) which are not included in (3.1) equal to an arbitrary definite quantity, so that expression (3.2) is well defined.

We restrict attention to hierarchical models, where whenever a particular u-term is zero, all of its higher-order relative must also be zero (e.g., if $u_{12(ij)} = 0$ for all pairs (i,j) for which it is defined in (3.1), then $u_{123(ijk)} = 0$ for all (i,j,k) ε S).

3.1 Unsaturated Models and Log-Linear Contrasts.

We define unsaturated quasi-log-linear models by setting u terms in (3.1) equal to zero, and the corresponding models can always be described in terms of generalized notions of interaction contrasts. For example, setting $u_{123(ijk)} = 0$ for all $(i,j,k) \epsilon S$ corresponds to setting equal to zero all generalized interaction contrasts of the form

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \delta_{ijk} \nu_{ijk} \log m_{ijk} \tag{3.6}$$

($\nu_{ijk} = 0$ for some i,j,k), where

$$\sum_{i=1}^{I} \delta_{ijk} \nu_{ijk} = \sum_{j=1}^{J} \delta_{ijk} \nu_{ijk} = \sum_{k=1}^{K} \delta_{ijk} \nu_{ijk} = 0. \tag{3.7}$$

Using (3.1) through (3.5), we can rewrite these three-factor interaction contrasts as

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \delta_{ijk} \nu_{ijk} u_{123(ijk)} \tag{3.8}$$

($\nu_{ijk} \neq 0$ for some i,j,k). We also have three different sets of two-factor interaction contrasts, which can be expressed in terms of linear contrasts of u-terms, such as

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \delta^{(3)}_{ij} \beta_{ij} u_{12(ij)} \tag{3.9}$$

($\beta_{ij} \neq 0$ for some i,j), where

$$\sum_{i=1}^{I} \delta^{(3)}_{ij} \beta_{ij} = \sum_{j=1}^{J} \delta^{(3)}_{ij} \beta_{ij} = 0 \tag{3.10}$$

Any unsaturated quasi-log-linear model can be specified by setting equal to zero appropriate sets of interaction contrasts, and a cell is noninteractive with respect to a quasi-log-linear model if it is not included in at least one of these interaction contrasts with nonzero coefficient.

3.2 Poisson likelihood function and Other Sampling Models.

We suppose that the observed count x_{ijk} in the (i,j,k) cell, for (i,j,k) ε S has a poisson distribution with mean m_{ijk}, the expected value for the (i,j,k) cell, and that these poisson variates are mutually independent. Then the kernel of the log-likelihood is

$$\Sigma_{ijk}\, x_{ijk} \log m_{ijk} = x_{+++}u + \Sigma_i\, x_{i++}\, u_{1(i)} + \Sigma_j\, x_{+j+}\, u_{2(j)}$$

$$+ \Sigma_k\, x_{++k}\, u_{3(k)} + \Sigma_{ij}\, x_{ij+}\, u_{12(ij)} + \Sigma_{ik}\, x_{i+k}\, u_{13(ik)}$$

$$+ \Sigma_{jk}\, x_{+jk}\, u_{23(jk)} + \Sigma_{ijk}\, x_{ijk}\, u_{123(ijk)}, \qquad (3.11)$$

since $m_{ijk} = 0$ implies $x_{ijk} = 0$. When, for example, $u_{123} = 0$ for all (i,j,k) ε S, then the sufficient statistics are once again given by the configurations $C_{12} = \{x_{ij+}\}$, $C_{23} = \{x_{+jk}\}$, and $C_{13} = \{x_{i+k}\}$.

Once again, as for complete multiway tables, we can use either a single multinomial sampling model or a set of multinomials based on exactly one fixed one-dimensional or two dimensional configuration, as long as the fixed configuration is included among the sufficient statistic.

4. SEPARABILITY IN MANY DIMENSIONS

We say that the (i,j,k) cell in an IxJxK table has first coordinate i, second j, and third coordinate k. Now if we let D be some nonempty subset of the set of integers $\{1,2,3\}$, we say that two cells are D-associated if they do not contain structural zeros and if their coordinates corresponding to the subset D coincide (e.g., if D = $\{1,3\}$ in a 2x4x2 table, then cells (1,3,2) and (1,4,2) are D-associated if they do not contain structural zeros). We say that a set of non-structural-zero cells is (D_1, D_2, D_3)-connected if any cell can be linked to any other cell via a chain of cells, any two consecutive members of which must be either D_1-associated, D_2-associated, or D_3-associated. Any set of non-structural-zero cells is (D_1, D_2, D_3)-separable if it is not (D_1, D_2, D_3)-connected.

We consider a set of non-structural-zero cells which is (D_1, D_2, D_3)-separable. We can divide this set up into subsets, each of which is itself (D_1, D_2, D_3)-connected but no two of which are (D_1, D_2, D_3)-connected whencombined. These subsets are referred to as the separable components of the original set.

Cohen [4] has suggested that the assessment of separability or inseparability of an incomplete multiway table is best done in two steps. The first step deals with the model, whether or not the full array includes structural zeros. Suppose we have a d-way table and that D = $\{1,\ldots,d\}$, the set of integers from 1 to d. We let $D_1, \ldots, D_m$ be the subsets of D which correspond to the highest-order u terms in the log-linear model. Then the complete d-way array is inseparable with respect to the model if and only if

$$\bigcup_{i=1}^{m} (D - D_i) = D.$$

Thus separability occurs when each of the subsets $\{D_i\}$ includes a common non-zero sub-subset. D*, i.e.,

$$\bigcap_{i=1}^{m} D_i = D^*.$$

More generally, for any separable model the complete array itself is separable, and it can be broken into subtables, each one of which corresponds to a cell in the cross-classification of the variables in D*. Then the log-linear model for the full table can be broken down into separate but parallel log-linear models for each subtable. In the assessment of separability for incomplete multiway tables, models for which the complete array is separable should be considered only in terms of the log-linear models for the subtables.

When an incomplete table is separable for a given quasi-log-linear model, then maximizing the likelihood function for that model turns out to be equivalent to maximizing separately the individual likelihood functions corresponding to each of the separable components or subtables. This is true because the maximum likelihood equations take the form of linear constraints on the expected cell values.

5. MAXIMUM LIKELIHOOD ESTIMATION

5.1 Conditions for Existence of MLEs.

We require that:

(a) The observed marginal configurations corresponding to the minimal sufficient statistics for the model must have positive entries whenever the corresponding expected configurations have positive entries.

(b) Both the expected and the observed table must be inseparable under the model.

When there exist unique nonzero MLEs for the nonzero expected cells of an incomplete multiway table and a particular quasi-log-linear model, these MLEs are uniquely determined by setting expected marginal configurations equal to the observed marginal configurations corresponding to the minimal sufficient statistics. If, for example, we are fitting the model specified by $u_{123} = 0$ in a three-way table, then the MLEs are given by the equations

$$\hat{m}_{ij+} = x_{ij+} \; ; \quad \hat{m}_{i+k} = x_{i+k} \; ; \quad \hat{m}_{+jk} = x_{+jk} \, , \tag{5.1}$$

where the subscripts in each set of equations range over all sets of values for which the expected marginal values are positive.

5.2 Iterative Procedure for Determining MLEs.

We can use the generalization of the Deming and Stephan [1] iterative Proportional Fitting (PF) procedure to compute estimated expected cell values.

1. At the 0th step, we let

$$m^{(0)}_{ijk} = \delta_{ijk} \tag{5.2}$$

for all i,j,k when δ_{ijk} is defined by (3.3) above.

2. At each successive cycle of the iteration, there are as many steps are there are configurations of sufficient statistics, and each step consists of a rescaling relative to óne of these configurations. For example, in a three-way table for $u_{123} = 0$, at the v^{th} cycle we take

$$m_{ijk}^{(3v-2)} = \frac{m_{ijk}^{(3v-3)} \quad x_{ij+}}{\Sigma_k \; m_{ijk}^{(3v-3)}}, \tag{5.3}$$

$$m_{ijk}^{(3v-1)} = \frac{m_{ijk}^{(3v-2)} \quad x_{+jk}}{\Sigma_k \; m_{ijk}^{(3v-2)}}, \tag{5.4}$$

$$m_{ijk}^{(3v)} = \frac{m_{ijk}^{(3v-1)} \quad x_{i+k}}{\Sigma_k \; m_{ijk}^{(3v-1)}}, \tag{5.5}$$

3. We continue repeating the cycles until desired accuracy is achieved.

This iterative procedure always converges, but the convergence tends to be somewhat slower than for the related procedure for complete multiway tables.

5.3 Degrees of Freedom.

To compute degrees of freedom for incomplete multiway tables, we subtract the number of independent parameters used in the model from the total number of cells to which the model is being fitted. This is somewhat more complicated since we may have structural zero totals in the expected marginal configurations associated with various models. Thus we must use the rule given by:

$$\text{degrees of freedom} = V - z_e + z_p \,, \tag{5.6}$$

where V is the number of degrees of freedom usually associated with the model for the complete table case, z_e is the number of cells with zero expected values, and z_p is the number of zero entries in the expected marginal configurations, adjusted for possible zeros in the marginal totals of these configurations.

We must look at each separable component of a seperable table by itself in order to compute the degrees of freedom properly for a given model as applied to the table as a whole. This is because separability leads to additional linear constraints on the expected cell values which are equivalent to our fitting extra parameters in the model. These constraints are "additional" in the sense that they do not result directly from the inclusion of particular parameters in the model for a general incomplete table.

To compute degrees of freedom with an IxJxK incomplete inseparable table for $u_{123} = 0$, containing a total of z_e structural zeros, we let

$$Z_{12} = IJ - \Sigma_{ij}\ \delta_{ij}^{(3)}$$

$$Z_{23} = JK - \Sigma_{jk}\ \delta_{jk}^{(1)} \tag{5.7}$$

$$Z_{13} = IK - \Sigma_{ik}\ \delta_{ik}^{(2)}$$

i.e. z_{12}, z_{23}, and z_{13} are the number of zeros in the expected marginal configurations $C^{*}_{12} = \{m_{ij+}\}$, $C^{*}_{23} = \{m_{+jk}\}$, and $C^{*}_{13} = \{m_{i+k}\}$, respectively. Thus $z_p = z_{12} + z_{23} + z_{13}$. Furthermore we let

$$\Sigma_i \ \delta_i^{(23)} = I, \quad \Sigma_j \ \delta_j^{(13)} = J, \quad \Sigma_k \ \delta_k^{(12)} = K \qquad (5.8)$$

Then the number of degrees of freedom is:

$$(I-1)(J-1)(K-1) - z_e + (z_{12} + z_{23} + z_{13}). \qquad (5.9)$$

If the same table is inseparable for the model with $u_{12} = u_{123} = 0$, the number of degrees of freedom for the latter model is then:

$$K(I-1)(J-1) - z_e + (z_{23} + z_{13}). \qquad (5.10)$$

If there is one empty layer in the table and (5.8) no longer holds, then further adjustments in the degrees of freedom formulas (5.9) and (5.10) are necessary. We must add one degree of freedom to (5.9) and (5.10). On the other hand, if there is one empty row in the table, we add one degree of freedom to (5.9) but not to (5.10).

6. AN APPLICATION-DISABILITY SCORES

The data in table 6.1 are presented originally in Bishop and Fienberg [2] and collected by Jones and Poskanzer at

Massachusetts General Hospital on 121 hospital patients. On admission and on discharge the patients were graded on a five-point scale (A through E) of increasing severity according to their physical disability following a stroke. Since no patient was discharged if he had become worse (except by death), a patient's score on the second examination could only be the same or better than on the original examination. The ordering of the disability scores, combined with this restriction, produces the block-triangular form of the table and form a two-way margin of multiway table. A third dimension of great interest sorted the patients on the basis of whether the stroke corresponded to a right (R) or left (L) lesion of the brain. This version of the data is given in table 6.2a, for which the rating scale has only three categories, I, II, and III, corresponding to the origianl categories in the following way:

$$A \ \& \ B \Longleftrightarrow I,$$
$$C \Longleftrightarrow II,$$
$$D \ \& \ E \Longleftrightarrow III.$$

The data have collapsed in this way on the basis of the original definition of disability categories as a result of the sparseness of data in parts of the table.

We begin the analysis by looking at quasi independence within each layer (i.e., for each of the two types of lesion). This can be done using closed-form methods for triangular tables, either separately for each type of lesion or by iteration, when we note that the conditional quasi-independence model is given by

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{13(ik)} + u_{23(jk)} ,$$

i.e,

[1] $u_{12} = u_{123} = 0.$

Table 6.1 Initial and Final; Ratings on Disability of Stroke Patients

	Final State					
Initial State	A	B	C	D	E	Totals
E	11	23	12	15	8	69
D	9	10	4	1	-	24
C	6	4	4	-	-	14
B	4	5	-	-	-	9
A	5	-	-	-	-	5
Totals	35	42	20	16	8	121

Source: Bishop and Fienberg [2].

Table 6.2 Three-Way Version of Jones and Poskanzer Data on Stroke Patients

a. Observed Data

	R-lesion				L-lesion			
	Final State				Final State			
Initial State	I	II	III	Totals	I	II	III	Totals
III	17	10	13	40	36	5	10	51
II	7	3	-	10	3	1	-	4
I	6	-	-	6	8		-	8
Totals	30	13	13	56	47	6	10	63

b. Expected Values for $u_{12} = u_{123} = 0$

	R-lesion				L-lesion			
	Final State				Final State			
Initial State	I	II	III	Totals	I	II	III	Total
III	17.51	9.49	13	40.00	35.53	5.47	10	51.0
II	6.49	3.51	-	10.00	3.47	0.53	-	4.00
I	6	-	-	6.00	8	-	-	8.00
Totals	30.00	13.00	13.00	56.00	48.00	7.00	10.00	63.0

c. Expected values for $u_{12} = u_{13} = u_{123} = 0$

	R-lesion				L-lesion			
	Final State				Final State			
Initial State	I	II	III	Totals	I	II	III	Total
III	20.35	10.78	13	44.13	31.89	4.98	10	46.87
II	4.19	2.22	-	6.41	6.57	1.02	-	7.59
I	5.46	-	-	5.46	8.54	-	-	8.54
Totals	30.00	13.00	13.00	56.00	47.00	6.00	10.00	63.0

$\therefore$ Expected Values for $u_{12} = u_{13} = u_{23} = u_{123} = 0$ (Complete Quasi Independence of Initial State, Final State, and Side of Lesion)

	R lesion				L lesion			
	Final State				Final State			
Initial State	I	II	III	Totals	I	II	III	Totals
III	24.59	7.41	10.82	42.82	27.66	8.34	12.18	48.18
II	5.06	1.53	-	6.59	5.69	1.72	-	7.41
I	6.59	-	-	6.59	7.41	-	-	7.41
Totals	36.14	8.94	10.82	56.00	40.76	10.06	12.18	63.00

The direct fitting is simple because the cells corresponding to the states (I,I) and (III,III) are cell isolates, and upon their deletion we look at the remaining 2x2 table in each layer. There are clearly two degrees of freedom for this model and table (one for each layer), and an examination of the observed and expected values shows an excellent fit, with $G^2 = 0.6$.

We can look at two quasi-log-linear models:

[2] $u_{12} = u_{13} = u_{123} = 0$,

[3] $u_{12} = u_{13} = u_{23} = u_{123} = 0$.

The expected values for these models, which can also be written in closed form, are given in tables 6.2c and 6.2d. The corresponding goodness-of-fit statistics are $G^2 = 5.49$ with four degrees of freedom and $G^2 = 11.89$ with six degrees of freedom.

respectively, neither of which is significant at the 5% level.

We note that models [1], [2], and [3] form a nested sequence with [1] as a special case of [2] and [2] a special case of [3] - for such nested sequences we can partition the likelihood ratio goodness-of-fit statistic G^2 into additive components. Thus the statistic for testing the fit of model [3] can be broken into two parts, one for testing the fit of [2], and the other for testing whether [3] is the true model given that [2] is also true. This second components has a value of 11.89 5.49 = 6.4 with two degrees of freedom, which is significant at the 5% level, even though the value of G^2 for model [3] is not. We can break the statistic for testing the fit of model [2] also into two parts, one for testing the fit of [1], and the other for testing whether [2] fits given that [1] does. The value of the latter component is 5.49-0.66 = 4.89 with two degrees of freedom, which is not significant the 5% level.

To summarize, we have a nested sequence of three models, each of which fits the data reasonably well. The difference between [1] and [2] is not significant at the 5% level, and so we opt for model [2] as opposed to [1] because of its simpler form. On the other hand, the difference between [2] and [3] is significant, and this leads us to choose [2] over [3].

We can therefore conclude that model[2], with $u_{12} = u_{13} = u_{12}$ = 0, is the appropriate one for the data at hand, and we can say that initial state is quasi independent of final state and side of lesion jointy. Since the rules for collapsing are applicable in this situation for collapsing over the categories for side of lesion, we see once again that initial state and final state are quasi-independent.

7. CONCLUSION

In this paper, we have discussed quasi-log-linear models for incomplete IxJxK three-way array, while attention has been restricted to hierarchical models. The model of quasi independence is linked to the definitions of separability and connectivity that depend in multiway tables both on the dimension of the table and on the particular model being fitted. The generalization for three-way tables is illustrated. It is shown that in the assessment of separability for incomplete multiway tables, models for which the complete array is separable should be considered only in terms of the log-linear models for the subtables.

Conditional quasi independence model is fitted to the data on stroke patients and compared with the fitting of two simpler quasi log-linear models. An examination of the observed and expected values using proportional fitting algorithm for the three models shows an excellent fit for model[1] while neither of the other two models is significant at the 5% level. Models [1], [2] and [3] form a nested sequence of three models, each of which fits the data reasonably well. Model[2] with $u_{12} = u_{13} = u_{123} = 0$ is the appropriate one for the data and we see that initial state is quasi-independent of final state and side of lesion jointly.

REFERENCES

[1] Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975), Discrete Multivaria Analysis: Theory and Practice, Cambridge, Mass.: *The MIT Press.*

[2] Bishop, Y. M. M. and Fienberg, S. E. (1989), "Incomplete two-dimensional conti gency tables," *Biometrics*, 25, pp. 119-128.

[3] Chen, W. Y., Crittenden, L. B., Manfel, N. and Cameron, W. R. (1961), "Si distribution of canser deaths in husband-wife and sibling pairs," *Journal o National Canser Institute*, 27, pp. 875-892.

[4] Cohen, J. F. (1973), "Ciledhood mortality, family size, and birth order in p lustrial Europe," *Applied Statistics*, 22, pp. 7-21.

[- [illegible]ing, W. E. and Stephen, F. F. (1940), "On a least squares adjustment [illegible]mpled frequency table when the expected marginal totals are knwon," *Annals o [illegible]hematics and Statistics*, 11, pp. 427-444.

[6] [illegible]enberg, S. E. (1969), "Preliminary graphical analysis and quasi-independence fo two-way contingency tables," *Applied Statistics*, 18, pp. 153-168.

[7] Fienberg, S. E. (1977), The Analysis of Cross-Classified Categorical Data, Camb ridge, Mass.: *The MIT Press.*

[8] Goodman, L.A. (1963), "Statistical methods for the preliminary analysis of tran saction flows," *Econometrika*, 31, pp. 197-208.

[9] Goodman, L.A. (1968), "The analysis of cross-classified data: independence, quasi independence, and interaction in contingency tables," *Biometrics*, 25, pp.119-128.

[10] Haberman, S. J. (1974), The Analysis of Frequency Data. Chicago, *University o Chicago Press.*

[11] Kastenbaum, M. A. (1958), "Estimation of relative frequencies of four sperm type in Drosophila melanogaster," *Biometrics*, 14, pp. 223-228.

[12] Mantel, N. and Halperin, M. (1963), "Analysis of birth-rank data," *Biometrics*, pp 875-892.

[13] Savage, I. R. and Deutsch, K. W. (1960), "A statistical model: the gross analysis o transaction flows," *Econometrika*, 28, pp. 551-572.

[14] Watson, G. S. (1959), "Some recent result in chi-square goodness-of-fit tests," *Biometrics*, 15, pp.440- 468.