Statistica & Applicazioni
Vol. II, n. 1, 2004
pp. 89-100

Models for contingency tables specified by linear constraints on two types of odds ratios

Manuela Cazzaro[§]

Summary:

A class of models for multi-way contingency tables obtained from a log-linear model by imposing non linear constraints on the interaction parameters are presented. In particular the constraints will regard measures of bivariate association between a pair of categorical variables based on the local-continuation and continuation-local odds ratios.
The contribution of this work, according to the author's knowledge, is the simultaneous constraining of more than one type of odds ratio and the approach to manage the row and the column effect of the dependence relationship between categorical variables.

Keywords: *Log-linear Models, Logit Models, Generalised Odds Ratios, Constrained Maximum Likelihood.*

1. Introduction

Several types of generalized odds ratios (Douglas-Fienberg-Lee-Sampson and Whitaker, 1990) can be defined by partitioning rectangular subarrays of a contingency table into four sets of adjacent cells. A variety of alternatives to independence can be obtained by constraining different types of odds ratios. An example is the situation, described by Schriever (1983) as double positive regression dependence of order one, where both the *local-global* and the *global-local odds ratios* are assumed to be greater or equal to one.

In this work a generalization of the *Goodman row effect and column effect association models*, (Goodman, 1979) is introduced by constraining the *local-continuation odds ratios* and the *continuation-local odds ratios*. The

[§] Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali – Università degli Studi di Milano-Bicocca – P.zza dell'Ateneo Nuovo, 1, 20127 MILANO (e-mail: manuela.cazzaro@unimib.it).

main difference between the proposed model and the Goodman model is on the type of odds ratios used: in the Goodman models the local odds ratios are constrained, in the generalization proposed in this work both the local-continuation and the continuation-local odds ratios are constrained.

The paper is organized as follows: section 2 examines the association constraints in the simple case of a two way contingency table in order to introduce the main ideas of the proposed approach; section 3 extends the analysis to the case of q categorical variables; section 4 deals with the maximum likelihood estimates of the parameters of the model proposed. Finally, in section 5 some applications are presented in order to exemplify the model.

2. The bivarate case

At first it will be examined the simple but important case of a two way contingency table. The general case of a multiway table is postponed to the next paragraph.

Let $\pi_{i_1 i_2}$ be the joint probabilities of two ordinal categorical variables A and B having respectively I_1 and I_2 categories. Let $\pi_{i_2/i_1}^{B/A}$ be the conditional probability of the category i_2 of B given the category i_1 of A. The continuation logit parameterization of the previous conditional probabilities is defined as follows:

$$\pi_{i_2/1}^{B/A} = \frac{exp\left(-\mu_{i_2}^{B}\right)}{\prod_{m=1}^{i_2}\left(1+exp\left\{-\mu_{m}^{B}\right\}\right)} \qquad i_2 < I_2 \tag{1}$$

$$\pi_{i_2/i_1}^{B/A} = \frac{exp\left(-\mu_{i_2}^{B} - \sum_{n=1}^{i_1-1}\varphi_{ni_2}\right)}{\prod_{m=1}^{i_2}\left(1+exp\left\{-\mu_{m}^{B} - \sum_{n=1}^{i_1-1}\varphi_{nm}\right\}\right)} \qquad i_1 > 1 \quad i_2 < I_2 . \tag{2}$$

The parameters $-\mu_{i_2}^{B}$ are continuation logits calculated on the first row and are defined as follows:

$$-\mu_{i_2}^{B} = ln\frac{\pi_{1i_2}}{\sum_{m=i_2+1}^{I_2}\pi_{1m}} \qquad i_2 = 1,2,...,I_2 - 1, \qquad \mu_{I_2}^{B} = 0.$$

The parameters

$$\varphi_{i_1 i_2} = ln \frac{\pi_{i_1 i_2} \cdot \sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}}{\pi_{i_1+1 i_2} \cdot \sum_{m=i_2+1}^{I_2} \pi_{i_1 m}}$$
$$i_1 = 1,2,...,I_1 - 1, \qquad i_2 = 1,2,...,I_2 - 1,$$

are the logarithms of the local-continuation odds ratios (o.r.) and describe the differences of the conditional distributions respect to the first row of the contingency table. From the previous definition it stands out that:

$$\frac{\pi_{i_1+1 i_2}}{\sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}} = exp\{-\varphi_{i_1 i_2}\} \frac{\pi_{i_1 i_2}}{\sum_{m=i_2+1}^{I_2} \pi_{i_1 m}}, \tag{3}$$

where $exp\{-\varphi_{i_1 i_2}\}$ is the factor that multiplies the row continuation odd $\pi_{i_1 i_2} \Big/ \sum_{m=i_2+1}^{I_2} \pi_{i_1 m}$ in order to obtain the odd $\pi_{i_1+1 i_2} \Big/ \sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}$ in the next row of the contingency table.

Relevant constraints on the local-continuation odds ratios that model the dependence of *B* from *A* in a parsimonious way are

- *Uniform dependence* of *B* from *A* or *row effect model*

$$\varphi_{i_1 i_2} = \varphi_{i_1 \bullet}. \tag{4}$$

  Under this model all the I_2-1 continuation odds of the conditional probability function of *B*, given the category i_1+1 of *A*, differ from the corresponding continuation odds of the conditional probability function of *B* given the category i_1 of *A* for the same proportionality factor.

- *Linear continuation logit regression* or *column effect model*

$$\varphi_{i_1 i_2} = \varphi_{\bullet i_2}.$$

  Under this model the following linear model holds for the continuation logits

$$ln \frac{\pi_{i_1+1 i_2}}{\sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}} = -\mu_{i_2}^{B} - \varphi_{\bullet i_2} \cdot i_1. \tag{5}$$

This model implies linear effect of the row index on the continuation logit of the row conditional distribution.

Formulae (4) and (5) are examples of linear constraints that model the dependence of *B* from *A*.

Reversing the role of *A* and *B* an analogous parameterization can be defined with *B* as the conditioning variable and *A* as the dependent one. Models similar to the previous ones, describing the dependence of *A* from *B* by involving the column continuation odds, can be easily obtained. In this context the local-continuation odds ratios are replaced by the parameters

$$\psi_{i_1 i_2} = ln \frac{\pi_{i_1 i_2} \cdot \sum_{n=i_1+1}^{I_1} \pi_{n i_2+1}}{\pi_{i_1 i_2+1} \cdot \sum_{n=i_1+1}^{I_1} \pi_{n i_2}}$$

$$i_1 = 1,2,...,I_1 - 1, \qquad i_2 = 1,2,...,I_2 - 1,$$

that are the logarithms of the so called *continuation-local o.r.*

With this kind of parameterization different constraints can be imposed on the continuation-local odds ratios as it was done for the models proposed for the local-continuation odds ratios.

In a contingency table both the dependence of *A* from *B* and the dependence of *B* from *A* are sometimes interesting rather than the dependence of one variable from the other one only (for example, the job satisfaction could depend on insomnia and viceversa). In these situations, constraints on the dependence of *A* from *B* and of *B* from *A* can be considered simultaneously. These considerations can be applied in those situations where in the two way contingency table are involved, for example, the same variable measured on two different occasions (classification of the days according to pollution level in two different areas, A_1 and A_2; classification of the right and left eye-functionality) and the same variable measured in two different units (classification of father's and son's political orientation).

There are different reasons that justify the simultaneous use of two asymmetrical types of o.r., the local-continuation and the continuation-local o.r. instead of the simplest symmetric o.r., the local one (as Goodman, 1979) for both conditional distributions. First of all in many situations it should be considered that it is more obvious and reasonable the use of asymmetrical odds ratios in order to measure the dependence in a contingency table. In these situations it is not convenient the use of symmetrical approach even if it will make the computational context easier. These are the cases where the dependence of *A* from *B* and of *B* from *A* are better described by continuation type odds ratios.

Hence constraints based on both types of odds ratios may simultaneously be imposed. The following hypotheses constrain both local-continuation and continuation-local odds ratios:

- *Double uniform dependence* (continuation-local column effect and local-continuation row effect)*:*

$$\varphi_{i_1 i_2} = \varphi_{i_1 \bullet}, \qquad \psi_{i_1 i_2} = \psi_{\bullet i_2}, \tag{6}$$

under this model there is uniform dependence of both *B* from *A* and *A* from *B*.

- *Double linear continuation logit regression* (continuation-local row effect and local-continuation column effect):

$$\varphi_{i_1 i_2} = \varphi_{\bullet i_2}, \qquad \psi_{i_1 i_2} = \psi_{i_1 \bullet}, \tag{7}$$

under this model the following linear continuation logits models hold:

$$ln \frac{\pi_{i_1+1 i_2}}{\sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}} = -\mu_{i_2}^{B} - \varphi_{\bullet i_2} \cdot i_1,$$

$$ln \frac{\pi_{i_1 i_2+1}}{\sum_{n=i_1+1}^{I_1} \pi_{n i_2+1}} = -\mu_{i_1}^{A} - \psi_{i_1 \bullet} \cdot i_2$$

where $\pi_{i_1+1 i_2} \Big/ \sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}$ denotes the continuation odds of the conditional probability function of *B* given *A* as previously established, and $\pi_{i_1 i_2+1} \Big/ \sum_{n=i_1+1}^{I_1} \pi_{n i_2+1}$ represents the continuation odds of the conditional probability function of *A* given *B*.

- *Continuation odds ratios symmetry model* (for square tables):

$$\varphi_{i_2 i_1} = \psi_{i_1 i_2}. \tag{8}$$

Under this model the factor that multiplies the row continuation odds $\pi_{i_1 i_2} \Big/ \sum_{m=i_2+1}^{I_2} \pi_{i_1 m}$ in order to obtain the odds in the next row $\pi_{i_1+1 i_2} \Big/ \sum_{m=i_2+1}^{I_2} \pi_{i_1+1 m}$ is the same to the one that gives the column

continuation odd $\pi_{i_1 i_2+1} \Big/ \sum_{n=i_1+1}^{I_1} \pi_{n i_2+1}$ from the odds in the previous column $\pi_{i_1 i_2} \Big/ \sum_{n=i_1+1}^{I_1} \pi_{n i_2}$.

It is worthwhile to observe that the constraints on the local-continuation odds ratios and the constraints on the continuation-local odds ratios treat the categorical variables in a non symmetric way. Anyway a certain form of symmetry is induced by the models (6)-(8) because they state analogous continuation logit models for the conditional distribution of *A* given *B* and of *B* given *A*. This sort of modelling strategy seems to be new and more flexible and it gives models easier to interpret than the usual strategies based on symmetric odds ratios as the local and global ones. A similar method based on the asymmetric local-global and global-local odds ratios is also possible and it has been used previously by Cazzaro and Colombi, 2000.

It is proposed to maximize the log-likelihood of the log-linear saturated model under all the constraints described previously in order to get the estimates of the previous parameters, $\mu_{i_1}^{A}$, $\varphi_{i_1 i_2}$, $\psi_{i_1 i_2}$ and $\mu_{i_2}^{B}$, under one of the proposed models that restricts both row conditional probabilities and column conditional probabilities and possibly also the marginal ones. This approach and its advantages will be discussed in paragraphs 4 and 5 after the introduction of the general case of a multiway table.

3. The general model

The starting point is the following log-linear model of constant partial association for the joint probabilities of q ordinal categorical variables A_s, s=1,…,q:

$$ln(\pi_{i_1 i_2 \ldots i_s \ldots i_q}) = \lambda_0 + \sum_{h=1}^{q} \lambda_{i_h}^{A_h} + \sum_{h=1}^{q-1} \sum_{k=h+1}^{q} \lambda_{i_h i_k}^{A_h A_k} \quad , \qquad (9)$$

where the $\lambda_{i_h}^{A_h}$ and $\lambda_{i_h i_k}^{A_h A_k}$ parameters satisfy the usual conditions of identifiability. If necessary the model could include more interactions even if the (9) is quite attractive as it is the simpliest model that allows to model the bivariate probabilities. In fact if the aim is to apply constraints on the bivariate probabilities in order to model these probabilities directly and on the conditional dependence hypotheses, a simpliest model than (9) is unreasonable. If the model (9) is not correctly specified, generally due to the absence of interactions of order greater than one, one could use a more complicated model in order to improve the fit.

The model (9) is used as a working model for the following reasons: a) under multinomial sampling the score functions of this model give a set of estimating equations that depend only on the bivariate marginal sample frequencies; b) the maximum likelihood estimators of the marginal bivariate probabilities are consistent even if the model (9) is not correctly specified.
Obviously the results presented for model (9) hold also for the saturated log-linear model of the joint probabilities $\pi_{i_1 i_2 \ldots i_s \ldots i_q}$.

The continuation-local and local-continuation odds ratios are used to model the bivariate association among the variables in the way described in the previous paragraph. By following this approach, constraints on the bivariate probabilities will be defined by constraining the following o.r.:

$$\varphi_{i_h i_k}^{A_h A_k} = ln \frac{\pi_{i_h i_k}^{A_h A_k} \cdot \sum_{m=i_k+1}^{I_k} \pi_{i_h+1 m}^{A_h A_k}}{\pi_{i_h+1 i_k}^{A_h A_k} \cdot \sum_{m=i_k+1}^{I_k} \pi_{i_h m}^{A_h A_k}} \tag{10}$$

$$\psi_{i_h i_k}^{A_h A_k} = ln \frac{\pi_{i_h i_k}^{A_h A_k} \cdot \sum_{n=i_h+1}^{I_h} \pi_{n i_k+1}^{A_h A_k}}{\pi_{i_h i_k+1}^{A_h A_k} \cdot \sum_{n=i_h+1}^{I_h} \pi_{n i_k+1}^{A_h A_k}} \tag{11}$$

$$i_h = 1,2,\ldots,I_h - 1, \qquad i_k = 1,2,\ldots,I_k - 1 .$$

Several models can be defined by constraining simultaneously the local-continuation and the continuation-local odds ratios like the *bivariate continuation odds ratios symmetry* hypothesis

$$\varphi_{i_k i_h}^{A_h A_k} = \psi_{i_h i_k}^{A_h A_k} \tag{12}$$

for square tables.

Other relevant models are the model of *bivariate double uniform dependence*

$$\varphi_{i_h i_k}^{A_h A_k} = \varphi_{i_h \bullet}^{A_h A_k}, \qquad \psi_{i_h i_k}^{A_h A_k} = \psi_{\bullet i_k}^{A_h A_k} \tag{13}$$

or the model of *bivariate double linear continuation logit regression*

$$\varphi_{i_h i_k}^{A_h A_k} = \varphi_{\bullet i_k}^{A_h A_k}, \qquad \psi_{i_h i_k}^{A_h A_k} = \psi_{i_h \bullet}^{A_h A_k} \tag{14}$$

It is worthwhile to note that the effects of covariates on bivariate probabilities can be easily modelled by allowing the odds ratios to be functions of covariates.

4. Maximum Likelihood estimates

In order to maximize the likelihood function of the model just presented under multinomial sampling, constraints should be considered. This problem of constrained optimization can be solved by using a generalization of the Aitchison-Silvey algorithm (Cazzaro-Colombi, 2002).

It is important to note that imposing constraints that restrict both row and column conditional probabilities implies to check some conditions that make the conditional distributions compatible for the existence of a joint probability distribution. In particular Arnold-Press (1989) showed that these compatibility conditions can be expressed constraining the tables containing the two conditional distributions, respectively, to have the same baseline odds ratios. This implies the equivalence between the local odds ratios computed on these tables. The algorithm used satisfies automatically the conditions of the compatibility constraints.

Let $\boldsymbol{\theta}$ be the parameter vector of the model (9) or of a more complicated model even the saturated one, and let us denote with $\boldsymbol{\theta}_0$ the vector of the ``true" values of parameters. Let $\mathbf{v}(\boldsymbol{\theta})$ be the vector of the linear constraints on the local-continuation odds ratios and on the continuation-local odds ratios.

It should be noted that $\mathbf{v}(\boldsymbol{\theta})$ contains linear constraints on the $\varphi_{i_k i_h}^{A_h A_k}$ and $\psi_{i_h i_k}^{A_h A_k}$ parameters but these constraints are non linear on the interaction parameters $\boldsymbol{\theta}$ of the log-linear model used.

The problem is to maximize the likelihood function $L(\boldsymbol{\theta})$ of the chosen log-linear model with respect to $\boldsymbol{\theta}$ under the constraints $\mathbf{v}(\boldsymbol{\theta}) = \mathbf{0}$.

Let $\mathbf{K} = \partial \mathbf{v}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}'$ be the matrix of the derivatives of these constraints with respect to $\boldsymbol{\theta}$.

If the model is correctly specified, from Aitchison-Silvey (1958) it follows that the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$, obtained under the constraints $\mathbf{v}(\boldsymbol{\theta}) = \mathbf{0}$, is such that $\sqrt{n}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)$ has an asymptotic normal distribution with a null vector of expected values and a variance matrix

$$\mathbf{V} = \left[\mathbf{F}^{-1} - \mathbf{F}^{-1}\mathbf{K}' \left[\mathbf{K}\mathbf{F}^{-1}\mathbf{K}'\right]^{-1} \mathbf{K}\mathbf{F}^{-1} \right]^{-1} \qquad (15)$$

where **F** is the average Fisher matrix of the log-linear model used. The matrices **V**, **F**, **K** in the previous statement are computed at $\boldsymbol{\theta}_0$. Obviously the matrices **K** and **F** differ by following the log-linear model chosen from the beginning.

The odds ratios $\varphi_{i_h i_k}^{A_h A_k}$, $\psi_{i_h i_k}^{A_h A_k}$ are functions of the joint probabilities which in turn are function of $\boldsymbol{\theta}$. From the previous remark it follows that a trivial application of the ``delta-method" and of the functional invariance property of the M.L. estimator gives the maximum likelihood estimators of the continuation odds ratios $\varphi_{i_h i_k}^{A_h A_k}$, $\psi_{i_h i_k}^{A_h A_k}$ and their asymptotic distribution.

5. Examples

In order to illustrate the models proposed in this work, the data reported in **Table 1** are analysed. These data concern 1164 medical service's users who are asked to evaluate their satisfaction (unsatisfied, satisfied, really satisfied) regarding the waiting-room's comfort (`Comfort') and the perception of the waiting-time in the waiting-room (`Time') before a specialist examination; they are also classified by sex and age (≤30, 31-50, 51-65, >65).

To begin a simple example will be presented where the hypotheses described in the second paragraph are tested only on one of the eight subtables of **Table 1**, then these hypotheses will be tested simultaneously on all the eight subtables. At first only the first subtable of the data will be considered: the two way contingency table (Time×Comfort) identified by the level `male' of the variable Sex and the level `≤30' of the variable Age.

Table 1. *Comfort, Time, Sex, Age*

SEX		male			Female		
COMFORT		uns.	sat.	r.sat.	uns.	sat.	r.sat.
AGE	**TIME**						
≤30	unsatisfied	5	5	2	14	10	4
	satisfied	3	10	3	17	26	11
	really sat.	2	5	9	3	17	32
31-50	unsatisfied	11	8	2	22	25	7
	satisfied	9	21	4	23	53	22
	really sat.	2	8	20	19	44	83
51-65	unsatisfied	8	6	2	16	10	4
	satisfied	8	29	9	12	28	23
	really sat.	2	18	37	8	24	48
>65	unsatisfied	6	9	4	2	6	9
	satisfied	5	29	14	7	28	22
	really sat.	4	16	47	1	28	74

Different models, imposing different hypotheses, have been fitted obtaining the following results:

1. continuation-local row effect: G^2=1.8351, df=2;
2. continuation-local column effect: G^2=1.3795, df=2;
3. local-continuation row effect: G^2=1.4473, df=2;
4. local-continuation column effect: G^2=1.8016, df=2;$
5. continuation-local row effect and local-continuation column effect (double linear continuation logit regression): G^2=2.6533, df=4;
6. continuation-local column effect and local-continuation row effect (double uniform dependence): G^2=1.9389, df=4;
7. continuation odds ratios symmetry model: G^2=5.6107, df=3.

All of these models fit the data quite well. The degrees of freedom of the models are given by the number of independent constraints imposed on the odds ratios.

The previous seven hypotheses can be tested also considering simultaneously all the 8 subtables of the dataset. In particular, these are the results:

1. continuation-local row effect: G^2=20.1351, df=16;
2. continuation-local column effect: G^2=18.1443, df=16;
3. local-continuation row effect: G^2=16.6950, df=16;
4. local-continuation column effect: G^2=29.3391, df=16;
5. continuation-local row effect and local-continuation column effect (double linear continuation logit regression): G^2=46.9530, df=32;
6. continuation-local column effect and local-continuation row effect (double uniform dependence): G^2=41.2682, df=32;
7. continuation odds ratios symmetry model: G^2=192.7424, df=24.

Most of these models fit the data quite well. Note as the model 6. is a good model and the model 5. is not. The degrees of freedom of these models are exactly eight times the ones of the hypotheses tested on just one subtable.

6. Conclusion and future developments

It has been shown in this paper how in a contingency table it is possible to model separately the dependence of a character *B* from a character *A* and the dependence of *A* from *B* by constraining two different types of odds ratios. It has been extended this approach to multiway tables in a straightforward manner. Moreover the properties of the M.L. estimator under the proposed constraints have been discussed and they have been allowed for a certain degree of mis-specification of the model.

When logarithms of local-continuation odds ratios and of continuation-local odds ratios, that in this work are subject to equality constraints, are all positive, both conditional distributions of *B* given *A* and the ones of *A* given *B* are ordered according to the criterium of uniform stochastic dominance.

This situation, using the local-global and the global-local odds ratios, has been discussed by Schriever (1983) under the name of *double order dependence of order one*. The hypothesis of positive local-continuation and continuation-local o.r. is relevant in many applications and it is important to be able to test this hypothesis along with the ones introduced in this work. This brings to the problem of testing simultaneously equality and inequality constraints. If the model (9) is correctly specified the results described in Colombi and Forcina (2000) and Fahrmeir and Klinger (1994) are easily extended in a straightforward way to the hypotheses of this work. However the problem of testing equality and inequality constraints, when the model is partially mis-specified, seems to be an open problem which will be studied deeply in another work.

Aknowledgments

The present work has been carried out within the following project: COFIN 2002, 2002133957_004.

References

Aitchison J., Silvey S. (1958). Maximum Likelihood Estimation of Parameters Subject to Restraints. *The Annals of Mathematical Statistics*, **29**, 813-828.

Arnold B.C., Press S.J. (1989). Compatible Conditional Distributions. *Journal of the American Statistical Association*, **84**, **405**, 152-156.

Cazzaro M., Colombi R. (2000). Parameters Estimation of a Multivariate Logit Model with Uniform Association Constraints. *Atti della XL Riunione Scientifica della Società Italiana di Statistica*, Firenze, 26-28 aprile 2000, 193-196.

Cazzaro M., Colombi R. (2002). A Hybrid Parameterization for Contingency Tables. In *Correlated data modeling*, D. Gregori, G. Carmeci, H. Friedl, A. Ferligoj, A. Wedlin, eds., 179-187, FrancoAngeli.

Colombi R., Forcina A. (2000). Modellizzazione di dati discreti con vincoli di uguaglianza e diseguaglianza. *Statistica*, **LX**, **2**, 195-214.

Douglas R., Fienberg S., Lee M., Sampson A., Whitaker L. (1990). Positive Dependence Concepts for Ordinal Contingency Tables, in *Topics in Statistical Dependence*, H. W. Block, A. R. Sampson and T. H. Savits eds., 189-202, Hayward, CA: Institute of Mathematical Statistics.

Fahrmeir L., Klinger J. (1994). Estimating and testing generalized linear models under inequality restrictions. *Statistical Papers*, **35**, 211-229.

Goodman L.A. (1979). Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories. *The Journal of the American Statistical Association*, **74**, 537-552.

Schriever B.F. (1983). Scaling of Order Dependent Categorical Variables with Correspondence Analysis. *International Statistical Review*, **51**, 225-238.