

Maximum likelihood estimators for log-linear models subject to constraints of double monotone dependence

Stimatori di massima verosimiglianza dei parametri di un modello log-lineare nell'ipotesi di doppia dipendenza monotona

Manuela Cazzaro

Dip. Met. Quant. per le Sc. Ec. Az.
Università di Milano-Bicocca
manuela.cazzaro@unimib.it

Roberto Colombi¹

Dip. Ing. Gest. e dell'Informazione
Università di Bergamo-Dalmine
colombi@unibg.it

Riassunto: La teoria asintotica degli stimatori di massima verosimiglianza in presenza di vincoli non lineari di disequaglianza dipende sia dalle usuali condizioni di regolarità (Birch (1964)) che giustificano l'approssimazione della log-verosimiglianza con una forma quadratica sia da una condizione sullo spazio parametrico definito dai vincoli di disequaglianza. Tale condizione è sostanzialmente riconducibile a una condizione di qualificazione dei vincoli del tipo di quelle usate nei problemi di ottimizzazione non-lineare. Scopo di questo lavoro è quello di esaminare l'importanza e l'utilizzo delle precedenti condizioni in un contesto di verifica di ipotesi sulla dipendenza di due mutabili. Il problema può essere considerato non standard sia per via della presenza di vincoli non lineari di disequaglianza sia per il fatto che questi sono in numero superiore al numero di parametri del modello log-lineare saturo per la funzione di probabilità multinomiale delle due mutabili.

Keywords: contingency tables, generalised odds ratio, ordered restricted inference.

1. Introduction

Let A and B be two qualitative variables with r and c categories. Sometimes both the dependence of A from B and the dependence of B from A are of interest. For example, the job satisfaction depends on insomnia and viceversa; the waiting-room's comfort influences the perception of the time while a patient is waiting in the waiting-room before a specialist examination and viceversa. These considerations are relevant also in those situations that involve, for example, the same variable measured on two different occasions or units (classification of the days according to the pollution level in two different areas, classification of the right and left's eye-functionality, classification of husband and wife's political orientation).

In these situations constraints on the dependence of A from B and B from A are of interest. To model the association between A and B usually a set of $(r - 1)(c - 1)$ symmetric odds ratios ($o.r.$), defined on the joint probabilities is subject to linear equality and inequality constraints. The most common types of symmetric odds ratios, that are constrained, are the local $o.r.$, the global $o.r.$ and the continuation $o.r.$

When the dependence of A from B and the dependence of B from A are of interest it seems more convenient to constrain two sets of asymmetric odds ratios defined respectively on the row conditional distributions of A given B and on the column condi-

¹Il lavoro è stato svolto nell'ambito del progetto COFIN 2002, protocolli 2002133957-002, 2002133957-004.

tional distributions of B given A . In this paper we will use the local-continuation and the continuation-local *o.r.* If π_{ij} are the joint probabilities, the logarithms of the local-continuation odds ratios are defined on adjacent rows as follows:

$$\varphi_{ij} = \ln \frac{\pi_{ij} \cdot \sum_{m=j+1}^c \pi_{i+1m}}{\pi_{i+1j} \cdot \sum_{m=j+1}^c \pi_{im}}, \quad i = 1, 2, \dots, r-1, \quad j = 1, 2, \dots, c-1$$

and the logarithms of the continuation local *o.r.* ψ_{ij} are analogously defined on adjacent columns of the contingency table. Alternatively the local-global and the global-local *o.r.* which are similarly defined can be used (for a survey of various type of odds-ratios see Douglas *et al.* (1990)). A context that makes worthwhile imposing constraints on two types of odds ratios simultaneously is the case in which we want to test that both sets of conditional distributions are stochastically ordered. For example the hypothesis $\varphi_{ij} \geq 0$ and $\psi_{ij} \geq 0$ of double positive dependence is equivalent to the hypothesis of uniform stochastic order of the row conditional distributions and of the column conditional distributions. A similar hypothesis of simple stochastic order is specified when the logarithms of the local-global and the global-local *o.r.* are assumed to be non-negative. A further relevant equality constraints hypothesis here considered, for square table, is the following new symmetry model: $\varphi_{ji} = \psi_{ij}$. Under this model the factors that multiply the continuation logits of the j -th row conditional distribution to obtain the corresponding logits in the next row are the same to the ones that give the continuation logits of the j -th column conditional distribution from the logits in the previous column. We are interested in testing the symmetry and double positive association hypothesis against the symmetry alternative and in testing for stochastic independence against the symmetry and double positive association hypothesis.

2. Constraints Qualification

The joint probabilities π_{ij} , of A and B , in lexicographical order, are the elements of the vector $\boldsymbol{\pi}$. These joint probabilities are supposed to be parameterized by the usual saturated log-linear model. The parameters vector of such a model will be represented by $\boldsymbol{\theta}$. The symmetry equality constraints and the double positive dependence inequality constraints will be denoted by

$$\mathbf{h}(\boldsymbol{\theta}) = \mathbf{0}, \mathbf{g}(\boldsymbol{\theta}) \geq \mathbf{0}.$$

Note that these equalities and inequalities are not linear in the vector $\boldsymbol{\theta}$ of log-linear parameters. For the details on how these constraints can be written see Colombi and Forcina (2001).

In addition to the parametric space $\Theta_H = \{\boldsymbol{\theta} : \mathbf{g}(\boldsymbol{\theta}) \geq \mathbf{0}, \mathbf{h}(\boldsymbol{\theta}) = \mathbf{0}\}$, defined by non linear equality and inequality constraints, we also introduce the parametric space $\Theta_0 = \{\boldsymbol{\theta} : \mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}, \mathbf{h}(\boldsymbol{\theta}) = \mathbf{0}\}$ obtained by replacing the inequality constraints by the equality ones. Later on $\mathbf{G}, \mathbf{H}$ will denote the jacobian matrices of $\mathbf{g}(\boldsymbol{\theta}), \mathbf{h}(\boldsymbol{\theta})$ respectively at the point $\boldsymbol{\theta}_0 \in \Theta_0$ which represents the unknown parameters vector for the model under the hypothesis that all the inequality constraints are satisfied like equalities.

It can be easily shown that $\mathbf{g}(\boldsymbol{\theta}), \mathbf{h}(\boldsymbol{\theta})$ satisfy the following Mangasarian-Fromovitz constraints qualification condition at the point $\boldsymbol{\theta}_0$:

$$\text{the set } C = \{\mathbf{d} : \mathbf{G}\mathbf{d} > \mathbf{0}, \mathbf{H}\mathbf{d} = \mathbf{0}\} \text{ is non empty.}$$

Note that the previous condition implies the Shapiro (1987) one:

$$clC = \{d : Gd \geq 0, Hd = 0\}$$

(where clC indicates the closure of C).

The conditions mentioned above are relevant in Nonlinear Programming to establish necessary optimality criteria (Mangasarian (1994)) and here, in the context of ordered restricted inference, are fundamental to obtain a reasonable asymptotic theory for the maximum likelihood estimators subject to inequality non linear constraints. Let $\hat{\theta}$, $\tilde{\theta}$, $\bar{\theta}$ be, respectively, the unrestricted maximum likelihood estimator, the one under $H_1 : \theta \in \Theta_H$ and the one under $H_0 : \theta \in \Theta_0$. Under the Mangasarian-Fromovitz condition and the usual regularity conditions the asymptotic chi-bar distribution for the test statistic $2(L_n(\tilde{\theta}) - L_n(\bar{\theta}))$ follows; an analogous result is valid for $2(L_n(\hat{\theta}) - L_n(\tilde{\theta}))$. It is important to stress that the chi-bar distribution, which is a mixture of chi-square random variables, is attained when inequalities occur only under the null hypothesis or only under the alternative hypothesis (Colombi and Forcina (2000)). For a detailed description of the numerical methods to compute the maximum likelihood estimators subject to inequality constraints and of the Montecarlo method to simulate the asymptotic p-values of the statistics $2(L_n(\tilde{\theta}) - L_n(\bar{\theta}))$, $2(L_n(\hat{\theta}) - L_n(\tilde{\theta}))$ see (Colombi and Forcina (2000), Colombi and Forcina (2001)). The importance of the Mangasarian-Fromovitz condition derives from the fact that it assures the local approximation of Θ_H at θ_0 with a polyhedral convex cone which in turn is needed to justify the asymptotic chi-bar distribution (Andrews (1999)).

3. An example

The models proposed in this work are used to analyse the data of Table 1, where medical service's users are asked to evaluate their satisfaction (unsatisfied, satisfied, really satisfied, extremely satisfied) regarding the waiting-room's comfort (C) and the perception of the waiting-time (T) in the waiting-room before a specialist examination.

Table 1. *The data.*

$\begin{matrix} C \\ T \end{matrix}$	<i>unsatisfied</i>	<i>satisfied</i>	<i>really satisfied</i>	<i>extremely satisfied</i>
<i>unsatisfied</i>	86	31	50	39
<i>satisfied</i>	34	40	42	34
<i>really satisfied</i>	53	50	99	76
<i>extr. satisfied</i>	43	44	121	367

Clearly the two variables are ordinal and we could expect a positive dependence in both direction: the satisfaction of the waiting-room's comfort influences the satisfaction respect to the perception of the waiting-time and viceversa. The following models were considered:

Model M1: double positive dependence model: $\varphi_{ij} \geq 0$ and $\psi_{ij} \geq 0$;

Model M2: continuation-local row effect model: $\psi_{ij} = \psi_{i\cdot}$;

Model M2a: double positive dependence and continuation-local row effect model: $\varphi_{ij} \geq 0$ and $\psi_{ij} = \psi_{i\cdot} \geq 0$;

Model M3: local-continuation column effect model: $\varphi_{ij} = \varphi_{\cdot j}$;

Model M3a: double positive dependence and local-continuation column effect model:

$\varphi_{ij} = \varphi_{.j} \geq 0$ and $\psi_{ij} \geq 0$;

Model M4: symmetry model: $\varphi_{ji} = \psi_{ij}$;

Model M4a: double positive dependence and symmetry model: $\varphi_{ji} = \psi_{ij} \geq 0$.

The results are summarised in Table 2, which indicates, when the likelihood ratio test statistic has an asymptotic chi-bar-squared distribution, the range of the degrees of freedom for the components of the mixture of chi-squared random variables.

Table 2. *Model results for the analysed data.*

<i>Model restriction</i>	<i>Test</i>	<i>Df</i>	<i>p-value</i>
<i>M1</i>	2.17052	0 – 13	0.92443
<i>M2</i>	32.8555	6	0.00000
<i>M2a-M2</i>	0.00	0 – 7	1
<i>M3</i>	49.0864	6	0.00000
<i>M3a-M3</i>	0.00	0 – 7	1
<i>M4</i>	12.3737	8	0.13529
<i>M4a-M4</i>	1.82534	0 – 7	0.83031

The assumption that both the conditional distribution respect the uniform stochastic order takes evidence from the results on model *M1*. Even the hypothesis of symmetry $\varphi_{ji} = \psi_{ij}$ may be reasonably accepted.

References

- Andrews D.W.K. (1999) Estimation when a parameter is on a boundary, *Econometrica*, 67, 1341–1383.
- Birch M.W. (1964) A new proof of the Pearson Fisher theorem, *Annals of Mathematical Statistics*, 35.
- Colombi R. and Forcina A. (2000) Modellizzazione di dati discreti con vincoli di uguaglianza e disuguaglianza, *Statistica*, 195–214.
- Colombi R. and Forcina A. (2001) Marginal regression models for the analysis of positive association of ordinal response variables, *Biometrika*, 88, 1007–1019.
- Douglas R., Fienberg S.E., Lee M.T., Sampson A.R. and Whitaker L.R. (1990) Positive Dependence Concepts for Ordinal Contingency Tables, in: *Topics in Statistical Dependence*, Block, H. W., Sampson, A. R. and Sanits T. H. editors, Institute of Mathematical Statistics, Lecture Notes, Monograph Series, Haywar, California.
- Mangasarian O.L. (1994) *Nonlinear Programming*, SIAM, Philadelphia.
- Shapiro A. (1987) On differentiability of metric projection, 1: Boundary case, *Proceedings of the American Mathematical Society*, 99.