

Testing inequality constraints in marginal models for bivariate Markov chains

Verifica di vincoli di disuguaglianza sui parametri dei modelli marginali per catene di Markov bivariate

Roberto Colombi¹ and Sabrina Giordano²

¹ Dip. di Ingegneria Gestionale e dell'Informazione, Università di Bergamo-Dalmine

² Dipartimento di Economia e Statistica, Università della Calabria

e-mail: sabrina.giordano@unical.it

Riassunto: Nell'ambito delle serie storiche di mutabili ordinali, descritte da catene di Markov bivariate del primo ordine, testare l'ipotesi che non esista dipendenza tra il passato di una mutabile ed il presente dell'altra (Granger noncausalità) si traduce in vincoli di uguaglianza su alcuni coefficienti nei modelli marginali, che parametrizzano le probabilità di transizione della catena doppia. All'ipotesi di noncausalità (Granger, 1969) vengono contrapposte opportune alternative espresse con vincoli di disuguaglianza sui parametri, che descrivono una certa dipendenza monotona di una mutabile dal passato dell'altra. Sotto condizioni di regolarità per catene di Markov e la linearità dei vincoli di disuguaglianza il test del rapporto delle verosimiglianze ha distribuzione asintotica chi quadro barrato il cui p -value può essere simulato. Tali ipotesi vengono testate per le serie storiche di dati sui servizi di rete wap e web di un'azienda di telefonia mobile.

Keywords: categorical time series, generalized linear model, stochastic ordering, order-restricted inference, chi-bar squared distribution.

1. Introduction

Let $\{A_t, B_t\}$ be a bivariate time series of ordinal categorical variables with a and b categories respectively. We assume that the joint behavior of the two series is described by a first order bivariate Markov chain (BMC_1). Thus, the set of $a \times b$ pairs of categories is the state space, whereas the probabilities $p_{ij|z_t}^{A_t B_t} = P[A_t = i, B_t = j | z_t]$, with $i = 1, \dots, a$ and $j = 1, \dots, b$, arranged into the vector π_t , define the joint transition probabilities at time t . The conditioning set z_t of the so-called "predetermined variables" contains two vectors, $I_{A_{t-1}}$ and $I_{B_{t-1}}$ of dummy variables standing for the realizations of A_{t-1}, B_{t-1} , and a vector of covariates x_t .

Note that the joint transition probabilities at time t may be thought of as joint probabilities in a two-way contingency table. This leads us to introduce a generalized linear model for the joint probabilities $p_{ij|z_t}^{A_t B_t} = P[A_t = i, B_t = j | z_t]$. In this regard, an invertible transformation $\eta_t = g(\pi_t)$ of the vector π_t of the joint transition probabilities satisfies the linear constraints $\eta_t = Z_t' \beta$, where the elements in Z_t are functions of the predetermined variables and β is a column vector of unknown parameters.

Under the first order markovianity assumption, the question that arises is whether the knowledge about B_{t-1} , given A_{t-1} , does not make forecasting the present A_t easier. The hypothesis is equivalent to stating that the conditional independence $A_t \perp B_{t-1} | A_{t-1}$ holds for all t . Under the BMC_1 assumption with positive transition probabilities, the

simple condition that has just been mentioned is equivalent to the general restriction $A_t \perp \mathbf{B}_{-\infty}^{t-1} | \mathbf{A}_{-\infty}^{t-1}$, known as Granger noncausality of $\{B_t\}$ on $\{A_t\}$, where the sets $\mathbf{A}_{-\infty}^{t-1} = \{A_{t-j}, j = 1, \dots, \infty\}$ and $\mathbf{B}_{-\infty}^{t-1} = \{B_{t-j}, j = 1, \dots, \infty\}$ contain the past path of the processes $\{A_t\}$ and $\{B_t\}$ (Colombi and Giordano, 2005).

It is worth noting that the main issue involved in the noncausality of $\{B_t\}$ on $\{A_t\}$, concerns only A_t , irrespective of the contemporaneous value of B_t at every time t . Parametric models which allow us to deal with such an issue are provided by the Marginal approach. In the Marginal modelling, π_t is reparameterized by marginal logits, defined on marginal transition probabilities $P[A_t = i | \mathbf{z}_t]$ and $P[B_t = j | \mathbf{z}_t]$, and by log odds ratios defined on joint transition probabilities. Thus the linear constraints are imposed on the vectors of marginal logits $\boldsymbol{\eta}_t^A$, $\boldsymbol{\eta}_t^B$ and on the vector of log odds ratios $\boldsymbol{\eta}_t^{AB}$. Under Marginal modelling it is possible to express the hypothesis concerning noncausality of $\{B_t\}$ on $\{A_t\}$ in the form of linear equality constraints on appropriate coefficients in the linear predictors which model the marginal logits of A_t .

To clarify, let's consider the linear predictors, with only additive effects, which model the marginal continuation logits of A_t , $\eta_{it}^A = \ln \frac{p_{i|\mathbf{z}_t}^{A_t}}{\sum_{l=i+1}^a p_{l|\mathbf{z}_t}^{A_t}}$, $i = 1, 2, \dots, a - 1$:

$$\eta_{it}^A = \beta_{i0}^A + \sum_{m=1}^{a-1} \beta_{im}^{AA} I_{\{A_{t-1}=m\}} + \sum_{n=1}^{b-1} \beta_{in}^{AB} I_{\{B_{t-1}=n\}} + \sum_{s=1}^K \beta_{is}^{Ax} x_t. \quad (1)$$

The parameters β_{in}^{AB} are baseline-continuation log odds ratios (baseline for B_{t-1} and continuation for A_t) and are the association measures between B_{t-1} and A_t at the baseline level of A_{t-1} (a and b are the baseline levels):

$$\beta_{in}^{AB} = \ln \frac{P[A_t = i | A_{t-1} = a, B_{t-1} = n] P[A_t > i | A_{t-1} = a, B_{t-1} = b]}{P[A_t > i | A_{t-1} = a, B_{t-1} = n] P[A_t = i | A_{t-1} = a, B_{t-1} = b]}.$$

Since the β_{in}^{AB} are log odds ratios, testing the hypothesis of noncausality $A_t \perp B_{t-1} | A_{t-1}$ for all t is equivalent to testing the linear equality constraints null hypothesis $H_0 : \beta_{in}^{AB} = 0$ for each i, n , against the alternative H_1 in which at least one β_{in}^{AB} is different from zero.

2. The order-restricted hypotheses

Beside the unrestricted alternative to the Granger noncausality hypothesis, order-restricted alternatives may also be of interest. More specifically we are interested in alternative hypotheses H_h that correspond to a form of monotone dependence of A_t on B_{t-1} . The nature of the dependence is specified by requiring that the conditional probabilities of A_t given B_{t-1} , satisfy an appropriate stochastic order for all values of A_{t-1} .

The stochastic order is chosen among the *simple stochastic*, *uniform* and *likelihood ratio* orderings. These three orderings are equivalent to the requirement that local-global, local-continuation and local-local log odds ratios respectively are nonnegative in all the two-way tables of the joint distribution of A_t and B_{t-1} given A_{t-1} (Bartolucci *et al.*, 2001).

For example with respect to the linear predictor (1), testing H_0 against:

$$\begin{aligned} H_h : \quad & \beta_{i1}^{AB} - \beta_{i2}^{AB} \geq 0, \beta_{i2}^{AB} - \beta_{i3}^{AB} \geq 0, \dots, \beta_{ib-1}^{AB} - \beta_{ib-2}^{AB} \geq 0, \beta_{ib-1}^{AB} \geq 0, \\ & i = 1, 2, \dots, a - 1, \end{aligned}$$

with at least one strict inequality is analogous to testing for Granger noncausality against uniform stochastic ordering of the conditional distributions of A_t given B_{t-1} for any given value of A_{t-1} .

The simple stochastic and the likelihood ratio orderings for these conditional distributions are the alternative hypotheses H_h to the G-noncausality hypothesis H_0 when the linear predictors (1) are specified in term of the global and of the local marginal logits of A_t respectively.

Another possibility is that inequality constraints may be present under the null hypothesis rather than the alternative one. Hence, if we suppose that some form of causality of $\{B_t\}$ on $\{A_t\}$ exists, the issue of interest becomes that of testing H_h against the unrestricted hypothesis H_1 already mentioned.

3. Likelihood ratio tests

Let L_0 , L_1 and L_h define the maximized log-likelihood functions under the hypotheses H_0 , H_1 and H_h previously introduced. Under suitable assumptions and in the case of the linear predictor (1), from Fahrmeir and Kaufmann, 1987, it follows that also in the context of nonstationary Markov chains the likelihood ratio test (LRT) $2(L_1 - L_0)$ for testing H_0 against H_1 has the classical chi-squared asymptotic distribution with $(a - 1)(b - 1)$ *df*. Instead, when the inequality constraints of monotone dependence define either the null hypothesis or the alternative, the asymptotic null distribution of the statistics $2(L_h - L_0)$ and $2(L_1 - L_h)$ for H_0 against H_h and H_h against H_1 are of the chi-bar squared type which, in this case, is a mixture of $(a - 1)(b - 1) + 1$ chi-squared random variables. This result follows from the same assumptions of Fahrmeir and Kaufmann, 1987, needed for the asymptotic distribution of $2(L_1 - L_0)$, and from the fact that the parametric space under H_h is defined by linear inequality constraints on the model parameters.

To test order-restricted hypotheses we need to compute the tail probability of the chi-bar squared distribution either by means of numerical computation of the weights or by estimating the *p-value* via Monte Carlo simulations (Sen and Silvapulle, 2005, Colombi and Forcina, 2001).

4. An example

We analyze a bivariate time series of the daily utilization rate level (low 0, medium 1, medium-high 2, high 3) of a server that provides connections for wap $\{A_t\}$ and web $\{B_t\}$ services for six months in a city of northern Italy.

The results of the LRT are summarized in the following table in which the hypotheses are expressed by means of four values; the first couple of values specifies the effect of the predetermined variables A_{t-1} , B_{t-1} on the marginal logits of A_t , the second couple of values does the analogous for the logits of B_t . The values 1, 0, 1+ correspond to the unrestricted effect, null effect (noncausality) and monotone dependence respectively.

Marginal models with additive effects as in the linear predictor (1), parallel continuation logits ($\beta_{im}^{AA} = \beta_{*m}^{AA}$, $i = 1, 2, \dots, a - 1$, $\beta_{in}^{AB} = \beta_{*n}^{AB}$, $i = 1, 2, \dots, a - 1$) and uniform association ($\eta_{ijt}^{AB} = \eta \forall i, j, t$) are considered.

Table 1: Tests on G-noncausality and order-restricted hypotheses in BMC_1

null hypothesis		alternative hypothesis		LRT	df	p-value
M_2	(1 1+, 1+ 1)	M_3	(1 1, 1 1)	0.6715	0 - 6	0.8930
M_1	(1 0, 0 1)	M_2	(1 1+, 1+ 1)	11.58482	0 - 6	0.0076
M_4	(1 0, 1+ 1)	M_5	(1 0, 1 1)	0	0 - 3	1
M_1	(1 0, 0 1)	M_4	(1 0, 1+ 1)	8.79662	0 - 3	0.0064
M_6	(1 1+, 0 1)	M_7	(1 1, 0 1)	0.03386	0 - 3	0.8754
M_1	(1 0, 0 1)	M_6	(1 1+, 0 1)	4.90272	0 - 3	0.05201

The double G-noncausality, of $\{A_t\}$ on $\{B_t\}$ and of $\{B_t\}$ on $\{A_t\}$, is rejected against the alternative hypothesis in which the conditional distributions of A_t given B_{t-1} respect the uniform ordering for any values of A_{t-1} and $\{A_t\}$ does not Granger cause $\{B_t\}$ ($M_1 vs M_6$). The Granger noncausality for both processes is also rejected when the uniform dependence between B_t and A_{t-1} , given B_{t-1} , and $\{A_t\}$ is not Granger caused by $\{B_t\}$, is the alternative ($M_1 vs M_4$), finally the same hypothesis is also rejected against the double uniform dependence hypothesis ($M_1 vs M_2$). Thus, the tests confirm that the knowledge of yesterday's requests for wap service has influence in determining the more probable level of the today's utilization of the server for the web service, given that the connections provided for the web the day before are known, and, viceversa, for the web on the wap. The double uniform monotonicity is accepted when compared with the unrestricted form of dependence ($M_2 vs M_3$). Moreover, by verifying $M_4 vs M_5$, the uniform dependence between B_t and A_{t-1} seems to hold also when the process $\{B_t\}$ does not Granger cause $\{A_t\}$. When the role of the two processes is inverted, the test $M_6 vs M_7$ leads to an analogous result.

References

- Bartolucci, F., Forcina, A., Dardanoni, V. (2001) Positive quadrant dependence and marginal modeling in two-way tables with ordered margins, *Journal of the American Statistical Association*, 1497–1505.
- Colombi, R., Forcina, A. (2001) Marginal regression models for the analysis of positive association of ordinal response variables, *Biometrika*, 1007–1019.
- Colombi, R., Giordano, S. (2005) Alcune indipendenze condizionali nelle serie storiche categoriali bivariate, *submitted to Statistica*.
- Fahrmeir, L., Kaufmann, H. (1987) Regression models for nonstationary categorical time series, *Journal of Time Series Analysis*, 8, 2, 147–160.
- Granger, C.W.J. (1969) Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, 37, 3, 424–438.
- Sen, P.K., Silvapulle, M.J. (2005) *Constrained Statistical Inference*, Wiley, New Jersey.