Likelihood-ratio tests for order-restricted log-linear models: A comparison of asymptotic and bootstrap methods

Francisca Galindo Garre [1]
Jeroen K. Vermunt
Marcel A. Croon
Tilburg University

Abstract

This paper discusses the testing of log-linear models with inequality constraints using both asymptotic and empirical approaches. Two types of likelihood ratio tests statistics are investigated: one comparing the order-restricted model to the independence model, which is the most restricted model, and the other comparing the order-restricted model to the saturated model, or the data. As far as the asymptotic approach is concerned, we will focus on the chi-bar-squared distribution and methods for obtaining the weights for this distribution. The proposed empirical approach makes use of parametric bootstrapping.

KEYWORDS: *log-linear models, inequality constraints, chi-bar-squared distribution, parametric bootstrap*

Resumen

TESTS DE RAZON DE VEROSIMILITUDES PARA HIPOTESIS DE ORDEN: UNA REVISIÓN DEL BOOTSTRAP PARAMETRICO Y DE LAS APROXIMACIONES PARA CALCULAR LOS PESOS DE LA DISTRUBUCION CHI CUADRADO BARRA

En este articulo se discute el ajuste de modelos loglineales con restricciones de desigualdad. Para ello se utilizan dos aproximaciones: una basada en la teor{ia asintótica y otra empírica. Dos tipos de estadísticos de razón de verosimilitudes son investigados: en el uno el modelo con restricciones ordinales se compara con el modelo de independencia, que es el modelo más restringido, en el otro el modelo con restricciones ordinales se compara con el modelo saturado que representa los datos. En cuanto a la teoria asintótica, la distribución chi cuadrado barra y los métodos para obtener los pesos en esta distribución son estudiados. La aproximación empírica se centra en el bootstrap paramétrico.

PALABRAS CLAVE: *Modelos loglineales, restricciones de desigualdad, distribución chi cuadrado barra, bootstrap paramétrico*

In the social sciences, the variables and the relationships studied often have an ordinal nature. Such ordinal variables can be analyzed in different manners. One option is to use methods for nominal variables, like simple log-linear models, which amounts to ignoring information about the order of categories. Other methods, like correspondence analysis and log-bilinear association models, estimate the unknown category scores of the ordinal variables. These models are, however, not strictly ordinal because the score parameters do not necessarily reflect the assumed direction of the association. A third approach involves estimating the model probabilities under specific inequality restrictions on relevant association measures. Such

---

[1] P.O.Box 90153 5000 LE Tilburg. The Netherlands: Department of Methodology and Statistics
E-Mail: f.galindogarre@uvt.nl

a nonparametric approach permits to define and test more intuitive hypotheses about ordinality. An example is a model that assumes that all local log-odds ratios are at least zero. Even though model estimation and testing is more complicated when adopting such a nonparametric approach, quite some work has been done on this topic (Robertson, Wright & Dykstra, 1988; Croon, 1990, 1991; Dardanoni & Forcina, 1998; Vermunt, 1999, 2001). An important reason why these nonparametric methods have not been extensively used so far is that this literature is not very accessible for applied researchers. One of the aims of this paper is to provide a less technical overview of this field.

Hypotheses involving maximum likelihood estimates are usually tested by means of the likelihood-ratio (LR) statistic. Under some regularity conditions, the LR statistic is asymptotically chi-squared distributed, where the number of degrees of freedom equals the difference between the number of free parameters in the two models that are compared to one another. This simple rule for obtaining the number of degrees of freedom can, however, not be applied when inequality constraints are imposed. The reason for this is that in such models the number of free model parameters depends on the sample. As a consequence, the asymptotic distribution of the LR statistic is no longer a unique chi-squared distribution, but a mixture of chi-squared distributions that is often referred to as chi-bar-squared.

The main difficulty of using asymptotic tests based on the chi-bar squared is the computation of the weights associated with the various numbers of degrees of freedom. Analytical solutions are only available if the number of inequality restrictions is smaller than 5. However, several methods to approximate the weights of the chi-bar-squared distribution have been developed. We will present the most important ones.

Rather than using an asymptotic approach to obtain the p value associated with the LR statistic, it can also be estimated using parametric bootstrapping. It is well-known that bootstrapping methods can be used to obtain an empirical approximation of the distribution of a test statistic when its asymptotic distribution is complicated or unknown (Langeheine, Pannekoek & Van de Pol, 1996). These methods have been successfully applied to the testing of various types of order-restricted models for categorical data. Ritov and Gilula (1993), for example, proposed such a procedure in maximum likelihood correspondence analysis with ordered category scores. Vermunt & Galindo (2001) showed that parametric bootstrapping offers reliable results when applied in order-restricted row-column association models. Vermunt (2001) used the method for testing order-restricted latent class models.

This paper gives a less technical overview of methods for the estimation and testing of log-linear models with inequality constraints. The next Section introduces the log-linear model with inequality constraints, describes maximum likelihood estimation by activated-constraints algorithms, and presents the relevant test statistics. Then we describe the asymptotic testing approach, discus the problems associated with the computation of the weights of the chi-bar-squared distribution, and explain the parametric bootstrapping method. Subsequently, the different approaches are compared with one another using an empirical example. The paper ends with a short discussion.

Log-linear models with inequality constraints

Log-linear definition of a positive association

Consider a log-linear model in which the logarithm of the expected frequency for data pattern i is given by,

$$\log m_i = \sum_{k=1}^{K} \beta_k X_{ik}, \tag{1}$$

with β_k denoting one of the K unknown parameters ($k = 1, ..., K$), and X_{ik} an element of the design matrix.

As is shown below, the hypotheses of a positive relationship can be tested by assuming that some of the two-variable term are at least zero. In other words, some of the parameters are restricted by the inequality constraint $\beta_k \geq 0$. This means that the K model parameters can be divided into two sets: a set of k_1 unrestricted parameters whose values can be any real number, and another set of k_2 order-restricted parameters, where $k_1 + k_2 = K$.

Consider the case of a 3-by-3 contingency table for which the independence model does not hold. The most natural manner to define the strength of the relationship between two variables in a log-linear analysis framework is via the local log-odds ratios, θ_{rc}, defined as

$$\log \theta_{rc} = \log m_{rc} + \log m_{r+1c+1} - \log m_{r+1c} - \log m_{rc+1}, \tag{2}$$

where r denotes a row, and c a column of the contingency table. If the two variables are ordinal and if their relationship is positive, one would expect each local log odds ratio to be non-negative. In other words, a positive relationship implies that

$$\log \theta_{rc} \geq 0.$$

By using an special coding scheme based on the differences between categories, it is possible to represent the $\log \theta_{rc}$ in terms of β_k parameters. In this system, each of the two-variable parameters corresponds to a local log-odds ratio. As a result, the constraint $\log \theta_{rc} \geq 0$ can be imposed via a log-linear model of the form (1) with constraints $\beta_k \geq 0$, for $k > k_1$.

[INSERT TABLE 1 HERE]

Table 1 gives the appropriate design matrix for the case of a 3-by-3 table. The first column containing only ones corresponds to the constant β_1. The next two columns represent the one-variable terms for the row variable. Because of the incremental coding, these two columns of the design matrix correspond to the difference between levels one and two and between levels two and three of the row variable, respectively. The same incremental coding is used for the column variable in the fourth and fifth column of the design matrix. Columns six to nine represent the two-variable interaction effects. As usual, these are obtained by multiplying the appropriate pairs of columns of the one-variable effects.

The logarithm of the expected frequency m_{rc} equals the scalar product of the row of the design matrix corresponding to pattern (r, c) and the vector of parameters. For example, $\log(m_{12})$ can be expressed as follows:

$$\log m_{12} = \beta_1 + \beta_2 + \beta_3 + \beta_5 + \beta_8 + \beta_9.$$

The correspondence between the log-odds ratios and two-way interaction parameters can easily be seen by replacing the logs of the expected frequencies appearing in equation (2) by the log-linear parameters. For example, the local log-odds ratio $\log \theta_{22}$ turns out to be equal to β_9; that is,

$$
\begin{aligned}
\log \theta_{22} &= \log(m_{22}) + \log(m_{33}) - \log(m_{32}) - \log(m_{23}) \\
&= (\beta_1 + \beta_3 + \beta_5 + \beta_9) + (\beta_1) - (\beta_1 + \beta_5) - (\beta_1 + \beta_3) \\
&= \beta_9.
\end{aligned}
$$

In the remaining of the paper, we will concentrate on this simple log-linear model for two-way tables. It should, however, be noted that the estimation and testing methods described can be used with any type of order-restricted log-linear model.

Maximum-likelihood estimation

An easy way to obtain maximum-likelihood (ML) estimates of the parameters of a model with inequality constraints is by means of an activated-constraints algorithm (Gill & Murray, 1974). An activated constraint is an equality restriction that is imposed (activated) when an inequality restriction is violated; in our case, it is an order-restricted parameter that is equated to zero if it would otherwise become negative. It is straightforward to convert the Newton-Raphson algorithm for standard log-linear models into an activated-constraints method.

In Newton-Raphson, the parameters are updated as follows:

$$\boldsymbol{\beta}^{(\nu)} = \boldsymbol{\beta}^{(\nu-1)} - (\mathbf{H}^{(\nu)})^{-1}\mathbf{q}^{(\nu)},$$

where ν represents the iteration number, $\mathbf{q}$ denotes the gradient vector containing the partial derivatives of the log-likelihood function with respect to the parameters to be estimated, and $\mathbf{H}$ denotes the matrix of the second partial derivatives, also called the Hessian matrix. In an activated-constraints variant of Newton-Raphson, the unrestricted and non-activated order-restricted parameters are updated in the usual manner. Parameters corresponding to activated constraints are only updated if the update is in the right direction; that is, if an update will yield a non-negative parameter value. This can be checked via the sign of the corresponding element of the gradient vector. After updating the parameters, it may be necessary to activate certain constraints; that is, if $\widehat{\beta}_k < 0$, for $k > k_1$. This procedure is repeated until some convergence criterion is reached.

Two likelihood-ratio tests

Let H_0 denote the model in which all k_2 order-restricted parameters are set equal to zero. In our case, H_0 equals to the independence model. moreover, let H_1 denote the order-restricted model and H_2 the model in which no restrictions are imposed on the log-linear parameters. In our case, H_2 is the saturated model. In order to test whether there is a positive relation between the two ordinal variables, we can either compare H_0 with H_1 or H_1 with H_2. Likelihood-ratio (LR) statistics are usually used for this purpose. The corresponding statistics, L_{01}^2 and L_{12}^2, are defined as

$$
\begin{aligned}
L_{01}^2 &= 2 \sum_i n_i \log \left(\frac{m_{i(1)}}{m_{i(0)}} \right) \\
L_{12}^2 &= 2 \sum_i n_i \log \left(\frac{m_{i(2)}}{m_{i(1)}} \right),
\end{aligned}
\tag{3}
$$

where n_i represents an observed frequency, and $m_{i(g)}$ an expected frequency under model g ($g = 0, 1, 2$). Both statistics measure discrepancies between two models: L_{01}^2 indicates whether the differences between the estimated frequencies under the independence model (H_0) and the ones under order-restricted model (H_1) are significant. If this is the case, there is evidence that we need the non-negative two-variable interaction terms to explain the data. L_{12}^2 tests whether the estimated frequencies under the order-restricted model (H_1) differ significantly from the data. If these differences are not significant, it can be concluded that the order-restricted model gives a good representation of the data.

An empirical example

The order-restricted log-linear model will be illustrated with an analysis of a two-way contingency table taken from Agresti's textbook "Categorical Data Analysis" (Agresti, 1990: Table 2.4). The two variables of interest are 'Income' and 'Job satisfaction'. Income is measured in dollars and has four levels. Job satisfaction also has four levels: very dissatisfied, little dissatisfied moderately satisfied, and very satisfied. The research question of interest is as to whether there is a positive relationship between income and job satisfaction.

<div align="center">[INSERT TABLES 2 AND 3 HERE]</div>

Table 3 reports the parameter estimates for the independence, the order-restricted, and the saturated model. If we look at the results obtained from the order-restricted model, we see that the constraints corresponding to the parameters β_{11} and β_{12} are activated while only one parameter, β_{11}, took a negative value in the saturated model. This illustrates that the constraints that should be activated to obtain the order-restricted ML solution cannot always be derived from the unrestricted model. The reason for this is that the parameters are not independent of one another.

In order to test H_0 versus H_1 and H_1 versus H_2, we should examine the values of L_{01}^2 and L_{12}^2. These take on the values 11.59 and 0.44, respectively, indicating that there is a

large discrepancy between the independence model and the order-restricted model and a small discrepancy between the order-restricted model and the data. The problem is, however, how to decide as to whether these discrepancies are significant. A naive approach to determine the p values corresponding to L_{01}^2 and L_{12}^2 would be to treat the activated constraints as a priori zeros and apply standard chi-squared tests. It this case, such a procedure would yield chi-squared tests with 2 and $(K - 2 - 1)$ degrees of freedom, respectively. Such a method is, however, incorrect because the number of activated constraints and, therefore, also the degrees of freedom depend on the sample. As is explained in more detail below, the appropriate method is to assume a chi-bar-squared distribution for L_{01}^2 and L_{12}^2.

The asymptotic method

The chi-bar-squared distribution

The likelihood-ratio test leads to rejection of the null hypothesis if the LR value exceeds the critical value corresponding to a nominal probability α, which is the maximum type I error that can be accepted. In order to find the critical value, we need the null asymptotic distribution of the statistic. The distribution for testing inequality constrains was first obtained by Bartholomew (1959), and subsequently studied by many other authors like Perlman (1969), Shapiro (1988), Wolak (1991), and Dardanoni & Forcina (1998).

By means of the Delta Method and the Central Limit Theorem, it can be shown that, under some regularity conditions, the LR statistic is asymptotically chi-squared distributed. One of these conditions is that the true parameter value is an interior point of the parameter space under the null hypothesis. With inequality restrictions, this condition need not be fulfilled because the true parameter value can be on the boundary of the parameter space. For this more general case, Shapiro (1985) showed that discrepancy statistics have the same asymptotic distribution as

$$\min_{y \in \Theta} (\widehat{\mathbf{y}} - \mathbf{y})' \mathbf{H}^{-1} (\widehat{\mathbf{y}} - \mathbf{y}),$$

where $\widehat{y}$ is a random variable with distribution $N(0, H)$, and Θ represents a cone (the part of the parameter space that is in agreement with the inequality constraints). Their asymptotic distribution is a chi-bar-squared $(\overline{\chi}^2)$ distribution, which is a mixture of chi-squared distributions given by

$$P\left[\overline{\chi}^2 \geq c\right] = \sum_{\ell=0}^{k_2} w_\ell\left(\mathbf{H}, \Theta\right) P\left[\chi_\ell^2 \geq c\right], \tag{4}$$

Here, χ_ℓ^2 denotes a chi-squared random variable with ℓ degrees of freedom for $\ell = 1, ..., k_2$, and $P\left[\chi_0^2 \geq c\right] = 0$. Furthermore, $w_\ell\left(\mathbf{H}, \Theta\right)$ denotes a non-negative weight that depend on the matrix $\mathbf{H}$ and on Θ. This weight represents the probability that exactly ℓ constraints are activated in a particular sample.

Let $\widehat{\beta}$ and β denote the maximum likelihood estimates and the true parameter values, respectively. It has been shown that, under some regularity conditions, the function $n^{\frac{1}{2}}\left(\widehat{\beta} - \beta\right)$ follows a multivariate normal distribution with mean zero and variance-covariance matrix H,

where H can be approximated by the Fisher information matrix. Since LR statistics measure the discrepancy between estimated parameters under two hypothetical models, and, following the results of Shapiro (1985), LR statistics are asymptotically $\bar{\chi}^2$ distributed.

The L_{01}^2 statistic measures the discrepancy between estimated parameters under the independence model (H_0) and the ones under the order-restricted model (H_1). It has the same distribution as the β that are in agreement with the ordering and that minimize the distance to the estimated parameters $\widehat{\beta}_0$ under the independence model. The matrix $\mathbf{H}$ can be replaced by the information matrix under the independence model, $\mathbf{H}_0$. That is,

$$\min_{\beta \in \Theta}(\widehat{\beta}_0 - \beta)'\mathbf{H}_0^{-1}(\widehat{\beta}_0 - \beta).$$

The asymptotic distribution is defined by equation (4).

The L_{12}^2 case is somewhat more complicated. This statistic represents the distance between the estimated parameters under the order-restricted model (H_1) and the ones under the saturated model (H_2). It has the same distribution as the β minimizing the distance with the estimated $\widehat{\beta}_1$ under the order-restricted model. Its asymptotic distribution equals

$$P\left[\bar{\chi}_{12}^2 \geq c\right] = \sum_{\ell=0}^{k_2} w_{k_2-\ell}\left(\mathbf{H}, \Theta\right) P\left[\chi_{k_2-\ell}^2 \geq c\right].$$

A problem in the choice of $\mathbf{H}$ arises from the fact that the number of activated constraints depends on the sample. As a result, the dimension of the vector of free parameters and the rank of the variance-covariance matrix vary from one sample to the other. Actually, the only way to find a critical value that does not depend on the number of activated constraints in a particular sample is by taking the least favorable case in which all the constraints are activated; that is $\mathbf{H} = \mathbf{H}_0$. As is shown below, this yield a somewhat conservative test.

The weights corresponding to $\bar{\chi}^2$

Because exact weights can only be calculated in certain special cases, several methods have been developed for approximating the weights of the chi-bar-squared distribution. The most important ones will be exposed in this section.

Direct calculation of weights

Robertson et al. (1988: section 2.4) and Shapiro (1985) showed that, under certain regularity conditions, weights can be calculated for $k_2 \leq 4$. For example, when $k_2 = 3$, the weights can be obtained as follows:

$$
\begin{aligned}
w_0(\mathbf{H}, \Theta) &= \frac{1}{4}\pi^{-1}\left(2\pi - [\cos(p_{12})]^{-1}[\cos(p_{13})]^{-1}[\cos(p_{23})]^{-1}\right), & (5) \\
w_1(\mathbf{H}, \Theta) &= \frac{1}{4}\pi^{-1}\left(3\pi - [\cos(p_{12.3})]^{-1}[\cos(p_{13.2})]^{-1}[\cos(p_{23.1})]^{-1}\right), \\
w_2(\mathbf{H}, \Theta) &= \frac{1}{2} - w_0(\mathbf{H}, \Theta), w_3(\mathbf{H}, \Theta) = \frac{1}{2} - w_1(\mathbf{H}, \Theta),
\end{aligned}
$$

where p_{ij} denotes element (i, j) of the matrix that is obtained by,

$$\mathbf{p} = (diag\mathbf{H}^{-1})^{-\frac{1}{2}}\mathbf{H}^{-1}(diag\mathbf{H}^{-1})^{-\frac{1}{2}},$$

and $p_{ij.k} = (p_{ij} - p_{ik}p_{ik})\left(1 - p_{ik}^2\right)^{-\frac{1}{2}}\left(1 - p_{jk}^2\right)^{-\frac{1}{2}}$ is the conditional correlation between elements i and j given k. Equation (5) gives an idea about the complexity of the computations for larger number of constraints.

Approximating the weights

Several methods have been developed to approximate the weights of the chi-bar-squared distribution when their values cannot be calculated directly. One of these methods consists of assuming that the information matrix $\mathbf{H}$ is the identity matrix ($\mathbf{I}$). Grove (1980) claimed that the $w_\ell(\mathbf{H}, \Theta)$ are insensitive to the choice of $\mathbf{H}$, and that, as a result, $w_\ell(\mathbf{I}, \Theta)$ provides a reasonable approximations in most situations.

Gourieroux et al. (1982) proposed approximating the weights of the chi-bar-squared by a Binomial distribution with k_2 trials and probability of success equal to $\frac{1}{2}$. In other words,

$$w_\ell(\mathbf{I}, \Theta) = 2^{-k_2}\frac{k_2!}{[\ell!\,(k_2 - \ell)!]},$$

where k_2 denotes again the number of order-restricted parameters.

Dardanoni & Forcina (1998) provided stochastic upper and lower bounds for the distribution of L_{01}^2 and a stochastic upper bound for the distribution of L_{12}^2 which depend on the type of order hypotheses. Here, we only give the bounds that apply to the model used in this paper. For L_{01}^2, these bounds have the following form:

$$\chi_1^2 \preceq_s \overline{\chi}_{01}^2 \preceq_s \overline{\chi}^2(\mathbf{I}_{k_2}, \Theta_{k_2}),$$

which means that for a certain critical value the cumulative probability under the asymptotic distribution of the statistic is contained in the interval determined by the cumulative probabilities under a chi-squared with one degree of freedom and a chi-bar-squared distribution defined in the restricted parameters space and having the identity matrix as a covariance matrix.

For L_{12}^2, the upper bound is given by

$$\overline{\chi}_{12}^2 \preceq_s \chi_{k_2-1}^2 + \chi_1^2,$$

which indicates that the cumulative probability under the asymptotic distribution of the statistic is smaller than the combination of the cumulative probabilities of chi-squared distributions with $k_2 - 1$ and one degree of freedom.

Estimating the weights by simulation

Dardanoni and Forcina (1998, p. 1117) proposed estimating the weights of the chi-bar squared distribution by means of a simulation procedure that makes use of the asymptotic distribution of the maximum likelihood estimators, $N\left(\widehat{\boldsymbol{\beta}}, \sqrt{n\mathbf{H}^{-1}}\right)$. Their procedure involves drawing a

reasonable number of parameter vectors from a normal distribution with mean equal to the hypothesized parameter values and a covariance matrix equal to the estimated information matrix under H_0. These simulated parameter vectors may contain values that violate the order restrictions. An activated-constraints algorithm is used to find order-restricted parameter values that are as close as possible to the simulated values in the weighted least squares sense. This procedure is sometimes referred to as projecting the simulated values into the restricted parameters space. The estimated weights of the chi-squared-bar are defined by the distribution of the number of activated constraints across replications.

In the case of the L_{01}^2 statistic, k_2 parameters are drawn from a normal distribution with mean equal to zero and covariance matrix equal to the information matrix under the independence model. The simulated parameters are projected into the space of non-negative parameters. Each weight is defined as the proportion of times that the corresponding number of activated constraints occurs in the replications. Using the simulated weights, the critical value can be obtained by equation (4).

The L_{12}^2 case is more complicated because the number of activated constraints in the model estimated under H_1 depends on the data. To circumvent this problem, Dardanoni and Forcina (1998) proposed using the least favorable case in which all the constraints are activated. This amounts to simulating the weights in the same manner as for the L_{01}^2 test. They also proposed an alternative, local, test in which the parameters are drawn from a multivariate normal distribution having the parameter estimates under the order-restricted hypothesis $\left(\widehat{\beta}_1\right)$ as mean, and the information matrix of that model as variance-covariance matrix. A disadvantage of this approach is that the approximation of the asymptotic distribution depends heavily on the number of activated constraints in the order-restricted model. An advantage is that it is less conservative.

The parametric bootstrapping method

For models as complex as the ones considered here, the parametric bootstrap seems to be an attractive method to obtain the p values associated with L_{01}^2 and L_{12}^2. The distribution of the test statistic is empirically reconstructed by drawing samples from the multinomial distribution defined by estimated probabilities under the more restricted model. This method has been used by various authors for testing models with inequality restrictions. For example, Ritov & Gilula (1993) proposed such a procedure in maximum likelihood correspondence analysis with ordered category scores, and Vermunt & Galindo (2001) applied the procedure in ordered row-column association models. Furthermore, Wang (1996) showed that critical values obtained by parametric bootstrapping are asymptotically consistent when testing stochastic ordering of several populations.

For L_{01}^2, the parametric bootstrap works as follows:

1. Estimate the model under H_0 and H_1, in this case, by using the activated-constraints algorithm.

2. Compute the test statistic.

3. Draw a sample of the same size as the original sample from the multinomial distribution defined by the probabilities under H_0.

4. Estimate the models defined by H_0 and H_1 with the generated sample and compute the test statistic $\left(L_{01}^2\right)^*$.

5. Repeat steps (3) and (4) a sufficiently large number of times B, yielding bootstrap replicates $\left(L_{01}^2\right)_1^*, ..., \left(L_{01}^2\right)_B^*$.

6. Use the empirical distribution of $\left(L_{01}^2\right)_1^*, ..., \left(L_{01}^2\right)_B^*$ to approximate the p-value by

$$\widehat{p}_B = \widehat{P}_B\left[\left(L_{01}^2\right)^* \geq c\right].$$

The estimated p-value is the fraction of bootstrap replicates $\left(L_{01}^2\right)^*$ exceeding the observed value of the test statistic for the given sample. The standard error of the estimated p-value equals $\sqrt{p\left(1-p\right)/B}$.

The bootstrap procedure for L_{12}^2 differs from the one for L_{01}^2 in that frequency tables have to be simulated from the estimated probabilities under H_1. Then, the order restricted model is estimated with the simulated samples, and the distance with the generated data is calculated by L_{12}^2. The p value and its standard error is computed as was described above for L_{01}^2 (see also, Vermunt, 1999).

Application of the testing methods in the empirical example

Let us return to the empirical example introduced in Section 2. In order to decide as to whether the discrepancies found in our example ($L_{01}^2 = 11.59$; $L_{12}^2 = 0.44$) are significant, we need to find either the associated p values or the critical values corresponding to a certain value of α. Note that since $k_2 > 4$, the weights of the corresponding chi-bar squared distribution must be approximated by one of the procedures described above.

[INSERT TABLE 4 HERE]

For a significance level of 0.05, Dardanoni and Forcina's (1998) method for obtaining bounds yields a critical value lying between 3.84 and 11.74 for L_{01}^2, and a critical value smaller than 19.3488 for L_{12}^2. Since $L_{01}^2 = 11.59$ and $L_{12}^2 = 0.44$, none of the models needs to be rejected when using this procedure.

With the parametric bootstrap method (1000 replications), we obtained p values of 0.0095 and 0.9125 for L_{01}^2 and L_{12}^2, respectively. According to this procedure, we would reject the independence model in favor of the order-restricted model and not reject the order-restricted model in favor of the saturated model. This indicates that there is evidence for a positive association between the two variables.

Table 4 reports chi-bar-squared weights approximated by several procedures, as well as the corresponding p values and critical values for $\alpha = 0.05$. As can be seen, we used binomial

weights and weights simulated from multivariate normal distributions under H_0 and H_1. For the latter, we used variants based on a Hessian matrix with dimension K and k_2, respectively. We will refer to these as $\mathbf{H}_1(K)$ and $\mathbf{H}_1(k_2)$.

As can be seen from Table 4, the critical values and the p values are strongly dependent on the method used to obtain the weights. For L_{01}^2, a procedure is more liberal, leads easier to rejection of the independence model in favor of the order-restricted model, if larger weights are given to the smaller numbers of activated constraints. In the case of L_{12}^2, the effect of the weights is the opposite: a procedure is more liberal if smaller weights are given to the smaller numbers of activated constraints. It should be noted that a more liberal procedure yields lower critical values and lower p values than a more conservative procedure.

For L_{01}^2, the binomial weights yield the most conservative test. According to this method, there is not enough evidence to reject the independence model. The conclusion is different if we use the simulated weights based on $\mathbf{H}_0$, in which case we reject the independence model in favor of the order-restricted model. Note that the latter procedure yields a p value that is very close to one obtained with the bootstrap method.

For L_{12}^2, the most liberal results come from the procedures using binomial weights and weights simulated using $\mathbf{H}_1(K)$. Note that the latter method gives a p value that is close to the one obtained with the parametric bootstrap. The other two procedures, simulating weights using $\mathbf{H}_1(k_2)$ or and $\mathbf{H}_0$, give almost the same results.

Discussion

Compared to standard log-linear models, the presented order-restricted models have the benefit that they permit more precise specification of the nature of the relationship between the variables of interest. Although for simplicity of exposition we concentrated on the analysis of two-way tables, the proposed approach can also be used with multi-way tables. As we saw, maximum likelihood estimation of log-linear models with inequality constraints is quite straightforward. Testing of such models is, however, somewhat more problematic because the results may dependent on the method that is used to obtain the critical value or the p value.

In the case of the L_{01}^2 test, simulating weights under H_0 and the bootstrap are the preferred methods. As could be expected, these two methods give very similar estimates of the p value. The bounds provided by Dardanoni and Forcina (1998) and the binomial weights yield too conservative tests. This is not a problem as long as the independence model is reject. If, however, as in our example, these procedures lead to acceptance of the independence model, it is wise to perform a bootstrap or simulate the weights of the chi-bar squared distribution in order to get a less conservative test.

In the L_{12}^2 case, the upper bound provided by Dardanoni and Forcina is too conservative and the binomial weights are much too liberal. Simulating weights using $\mathbf{H}_0$ or $\mathbf{H}_1(k_2)$ yield similar but somewhat conservative results. These procedures provide a kind of upper bound for the p value. The consequence of using such a too conservative upper bound is that the order-restricted model may be accepted even when the relationship between variables is very weak.

The procedures based on the bootstrap and on weights simulated with $\mathbf{H}_1(K)$, on the other hand, may yield somewhat too liberal results. This is confirmed by results we obtained with other data set to which we applied these methods. Parametric bootstrap and simulated weights under H_1 (with dimension K) are both affected by the number of constraints activated in the estimated model. The consequence of this dependence is that they seem to give a kind of lower bound for the p value.

Further research should be done on L_{12}^2 test along two lines. First, we wish to get more insight into the behavior of the too conservative procedures, simulating weights using $\mathbf{H}_0$ or $\mathbf{H}_1(k_2)$, and the too liberal procedures, bootstrapping and simulating weights using $\mathbf{H}_1(K)$. This involves performing an extended simulation study. A second line of research is the search for possible improvements of current procedures, as well as for other testing approaches, like Bayesian methods, that might solve the problems associated with the current methods. An example of a possible adaptation of the current procedures is the use of a double bootstrap to make it less dependent on the number of the activated constraints in the maximum likelihood solution.

References

Barlow, R.E., Bartholomew, J.M. Bremner, J.M., and Brunk, H.D. (1972) *Statistical Inference under Order Restrictions,* John Wiley and Sons, London.

Bartholomew, D.J. (1959). A test of homogeneity for ordered alternatives. *Biometrika,* 46,36-48.

Bartolucci, F. & Forcina, A. (2000) A likelihood ratio test for MTP_2 within binary variables. *The Annals of Statistics*, 28, 1206-1218.

Bartolucci, F., Forcina, A. & Dardanoni, V. (2001) Positive quadrant dependence and marginal modeling in two-way tables with ordered margins. *Journal of American Statistical Association*, 96, 1497-1505.

Colombi, R. & Forcina, A. (2001). Marginal regression models for the analysis of positive association of ordinal response variables. *Biometrika*, 88(4), 1007-1019.

Croon, M.A. (1990). Latent class analysis with ordered latent classes. *British Journal of Mathematical and Statistical Psychology, 43*, 171-192.

Croon, M.A. (1991) Investigating Mokken scalability of dichotomous items by means of ordinal latent class analysis. *British Journal of Mathematical and Statistical Psychology, 44*, 315-331.

Dardanoni, V. & Forcina A. (1998). A unified approach to likelihood inference on stochastic ordering in a nonparametric context. *Journal of the American Statistical Association, 93*, 1112-1123.

Geyer, C.J. (1991) Constrained maximum likelihood exemplified by isotonic convex logistic regression. *Journal of the American Statistical Association, 86*, 717-724.

Geyer, C.J. (1995) Likelihood ratio tests and inequality constraints. *University of Minnesota: Technical Report No. 610*

Gill, P.E., & Murray, W. (1974). *Numerical Methods for Constrained Optimization*. London: Academic Press Inc.

Grove, D.M. (1980). A test of independence against a class of ordered alternatives in a $2 \times C$ contingency table. *Journal of the American Statistical Association, 75*, 454-459.

Langeheine, R., Pannekoek, J., & Van de Pol, F. (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological Methods and Research, 24*, 492-516.

Perlman, M.D. (1969). One-sided testing problems in multivariate analysis. *Annals of Mathematical Statistics, 40*, 549-567.

Robertson, T., Wright, F.T., & Dykstra, R.L. (1988). *Order Restricted Statistical Inference,* New York: John Wiley.

Ritov, Y. & Gilula, Z. (1993) Analysis of contingency tables by correspondence models subject to order constraints. *Journal of the American Statistical Association,* 88, 1380-1387.

Shapiro, A. (1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints. *Biometrika,* 72, 133-144.

Shapiro, A. (1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *International Statistical Review,* 56, 49-62.

Shi, N. (1991). A test of homogeneity of odds ratios against order restrictions. *Journal of the American Statistical Association,* 86, 154-158.

Vermunt, J.K. (1999). Nonparametric models for ordinal data. *Sociological Methodology,* 29, 187-223.

Vermunt, J.K. (2001). The use of restricted latent class models for defining and testing nonparametric and parametric item response theory models. *Applied Psychological Measurement,* 25(3), 283-294.

Vermunt, J.K. & Galindo, F. (2001). RC association models with ordered row and column scores. *Submitted.*

Wang, Y. (1996). A likelihood ratio test against stochastic ordering in several populations. *Journal of the American Statistical Association,* 91, 1676-1683.

Wolak, F.A. (1991), The local nature of hypothesis tests involving inequality constraints in nonlinear models. *Econometrica,* 59, 981-995.

Table 1. Design matrix for a 3-by-3 table using difference coding

pattern i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	X_{i6}	X_{i7}	X_{i8}	X_{i9}
1 1	1	1	1	1	1	1	1	1	1
1 2	1	1	1	0	1	0	0	1	1
1 3	1	1	1	0	0	0	0	0	0
2 1	1	0	1	1	1	0	1	0	1
2 2	1	0	1	0	1	0	0	0	1
2 3	1	0	1	0	0	0	0	0	0
3 1	1	0	0	1	1	0	0	0	0
3 2	1	0	0	0	1	0	0	0	0
3 3	1	0	0	0	0	0	0	0	0

Table 2. Observed cross-classification of job satisfaction and income

	Job Satisfaction			
Income($)	Very dissatisfied	Little dissatisfied	Moderately satisfied	Very satisfied
≤6000	20	24	80	82
6000-15000	22	38	104	125
15000-25000	13	28	81	113
≥25000	7	18	54	92

Table 3. Parameter estimates of the three models estimated with the data of Table 2

Unrestricted	β_1	β_2	β_3	β_4	β_5	β_6	β_7		
H_0	4.36	-0.34	0.21	0.32	-0.56	-1.08	-0.26		
H_1	4.25	-0.42	0.10	0.21	-0.94	-1.10	-0.53		
H_2	4.52	-0.42	0.10	0.21	-0.94	-1.10	-0.53		
Restricted	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	β_{16}
H_1	0.22	0.26	0.19	**0.00**	**0.00**	0.02	0.11	0.16	0.20
H_2	0.36	0.22	0.18	**-0.20**	**0.06**	0.04	0.16	0.15	0.20

Table 4. Weights of the chi-bar-squared distribution and corresponding critical values and p values for the empirical example

#activated constraints	Binomial	Simulated H_0	Simulated $H_1(K)$	Simulated $H_1(k_2)$
0	0.0000	0.0474	0.0000	0.0000
1	0.0020	0.1780	0.0040	0.0140
2	0.0176	0.2896	0.0088	0.1067
3	0.0703	0.2631	0.0645	0.2583
4	0.1641	0.1515	0.2032	0.3065
5	0.2461	0.0554	0.3151	0.2072
6	0.2461	0.0132	0.2650	0.0849
7	0.1641	0.0018	0.1161	0.0201
8	0.0703	0.0001	0.0248	0.0022
9	0.0176	0.0000	0.0021	0.0002
$\overline{\chi}^2_{01;0.05}$	11.7376	7.7179		
$\overline{\chi}^2_{12;0.05}$	8.4903	13.7421	9.6466	11.6364
$P\left(\overline{\chi}^2_{01} \geq 11.59\right)$	0.0667	0.0099		
$P\left(\overline{\chi}^2_{12} \geq 0.44\right)$	0.8897	0.9960	0.9362	0.9822