Available online at www.sciencedirect.com

ScienceDirect

Applied Mathematics Letters 20 (2007) 7–12

Applied Mathematics Letters

www.elsevier.com/locate/aml

Conditional tests of marginal homogeneity based on ϕ-divergence test statistics

M.L. Menéndez[a,*], J.A. Pardo[b], L. Pardo[b], K. Zografos[c]

[a] Department of Applied Mathematics, Polytechnical University of Madrid, 28040 Madrid, Spain
[b] Department of Statistics and O.R. Complutense University of Madrid, 28040 Madrid, Spain
[c] Department of Mathematics, University of Ioannina, 45110 Ioannina, Greece

Received 25 March 2005; received in revised form 28 September 2005; accepted 9 February 2006

Abstract

In this work, using the well-known result that symmetry is equivalent to quasi-symmetry and marginal homogeneity simultaneously holding, two families of test statistics based on ϕ-divergence measures are introduced for testing conditional marginal homogeneity assuming that quasi-symmetry holds.

Keywords: Symmetry; Quasi-symmetry; Marginal homogeneity; ϕ-Divergence test statistics

1. Introduction

Let Ω be a population such that for each element $w \in \Omega$ we consider two discrete random variables X and Y taking the values $x_1, \ldots, x_I$ and $y_1, \ldots, y_I$, respectively. We define $p_{ij} = P(X = x_i, Y = y_j) > 0$, $i, j = 1, \ldots, I$. We consider from the population Ω a random sample of size n and define $N_{ij} = \sum_{l=1}^{n} I_{\{x_i, y_j\}}(X_l, Y_l)$, $i, j = 1, \ldots, I$. It is well known that the random variable $(N_{11}, \ldots, N_{II})$ is obviously sufficient for the statistical model under consideration and is multinomially distributed with parameters n and $(p_{11}, \ldots, p_{II})^T$. We also define $\hat{p}_{ij} = N_{ij}/n$ and denote by $\hat{\boldsymbol{p}} = (\hat{p}_{11}, \ldots, \hat{p}_{II})^T$ the vector of relative frequencies. We consider the parameter space

$$\Theta = \{\boldsymbol{\theta} : \boldsymbol{\theta} = (p_{ij}; i = 1, \ldots, I, j = 1, \ldots, I, (i, j) \neq (I, I))^T\} \quad (1)$$

and we denote by $\boldsymbol{p}(\boldsymbol{\theta}) = (p_{11}, \ldots, p_{II})^T = \boldsymbol{p}$ the probability vector characterizing our model with $p_{II} = 1 - \sum_{i=1}^{I} \sum_{\substack{j=1 \\ (i,j)\neq(I,I)}}^{I} p_{ij}$. With this notation the problems of Symmetry, Marginal Homogeneity and Quasi-symmetry can be characterized by

$$H_0 : p_{ij} = p_{ji}, \quad i, j = 1, \ldots, I, \quad (2)$$

* Corresponding address: Escuela tecnica Superior de Arquitectura, Matematica Aplicada, Avda Juan de Herrera n4, 28040 Madrid, Spain. Tel.: +34 913366562; fax: +34 913366563.
E-mail address: ml.menendez@upm.es (M.L. Menéndez).

0893-9659/$ - see front matter
doi:10.1016/j.aml.2006.02.015

$$H_0 : \sum_{i=1}^{I} p_{ji} = \sum_{i=1}^{I} p_{ij}, \quad j = 1, \ldots, I-1 \tag{3}$$

and

$$H_0 : p_{ij} p_{jI} p_{Ii} - p_{iI} p_{Ij} p_{ji} = 0, \quad i, j = 1, \ldots, I-1, \tag{4}$$

respectively.

The problem of symmetry was first discussed by Bowker [9] who gave the maximum likelihood estimator as well as a large sample chi-squared type test for the null hypothesis of symmetry. In [15] a minimum discrimination information estimator was proposed and in [24] a minimum chi-squared estimator. On the basis of the maximum likelihood estimator and on the family of ϕ-divergence measures, in [20] a new family of test statistics was introduced. This family contains as a particular case the test statistic given by [9] as well as the likelihood ratio test. The state-of-the-art in relation to the symmetry problem can be seen in [8,2,4,25] and references therein. The problem of marginal homogeneity was first discussed by Stuart (in 1955), who defined a test statistic which is a quadratic form in the differences of the corresponding marginal values, whose matrix is the inverse of a consistent estimate of the covariance matrix of the differences under the null hypothesis, and its asymptotic distribution is chi squared with $I - 1$ degrees of freedom under the null hypothesis of marginal homogeneity. This hypothesis has been discussed by several authors (e.g. [5,6,15,8,1,7,17]). Finally, the hypothesis of quasi-symmetry was introduced by Caussinus [10] who gave a maximum likelihood estimator for quasi-symmetry as well as a chi-squared type statistic for the test of this hypothesis. For additional discussion of quasi-symmetry, see [12,13,19,14,2,25]. Recently, Matthews and Crowther [18] studied quasi-symmetry and independence for cross-classified data in a two-way contingency table.

It is well known that the maximum likelihood estimators, $\hat{\boldsymbol{\theta}}^{\mathrm{S}}$ (Symmetry), $\hat{\boldsymbol{\theta}}^{\mathrm{MH}}$ (Marginal Homogeneity) and $\hat{\boldsymbol{\theta}}^{\mathrm{QS}}$ (Quasi-symmetry) are given by

$$D_{\mathrm{Kull}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S}})) = \inf_{\{\boldsymbol{\theta} \in \Theta : p_{ij} - p_{ji} = 0, i<j, i,j=1,\ldots,I\}} D_{\mathrm{Kull}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\theta})) \tag{5}$$

$$D_{\mathrm{Kull}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{MH}})) = \inf_{\{\boldsymbol{\theta} \in \Theta : \sum_{i=1}^{I} p_{ji} - \sum_{i=1}^{I} p_{ij} = 0, j=1,\ldots,I-1\}} D_{\mathrm{Kull}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\theta})) \tag{6}$$

$$D_{\mathrm{Kull}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS}})) = \inf_{\{\boldsymbol{\theta} \in \Theta : p_{ij} p_{jI} p_{Ii} - p_{iI} p_{Ij} p_{ji} = 0, i,j=1,\ldots,I-1\}} D_{\mathrm{Kull}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\theta})), \tag{7}$$

where $D_{\mathrm{Kull}}(\boldsymbol{p}\boldsymbol{q})$ is the Kullback–Leibler measure of divergence, see [16], between the probability vectors $\boldsymbol{p} = (p_{11}, \ldots, p_{II})^{\mathrm{T}}$ and $\boldsymbol{q} = (q_{11}, \ldots, q_{II})^{\mathrm{T}}$, defined by

$$D_{\mathrm{Kull}}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^{I} \sum_{j=1}^{I} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{8}$$

In [21] the three problems were studied using the restricted minimum ϕ-divergence estimator. This estimator is based on the ϕ-divergence measure defined independently by [11] and [3],

$$D_{\phi}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^{I} \sum_{j=1}^{I} q_{ij} \phi\left(\frac{p_{ij}}{q_{ij}}\right); \quad \phi \in \Phi^* \tag{9}$$

where Φ^* is the class of all convex functions $\phi(x)$, $x > 0$, such that at $x = 1$, $\phi(1) = \phi'(1) = 0$, $\phi''(1) > 0$, and at $x = 0$, $0\phi(0/0) = 0$ and $0\phi(p/0) = \lim_{u\to\infty} \phi(u)/u$. For more details about ϕ-divergence measures, see [23].

The restricted minimum ϕ-divergence estimators for the problems considered in (2)–(4) could be obtained as the values $\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}$ (Symmetry), $\hat{\boldsymbol{\theta}}^{\mathrm{MH},\phi}$ (Marginal Homogeneity), and $\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}$ (Quasi-symmetry) verifying

$$D_{\phi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi})) = \inf_{\{\boldsymbol{\theta} \in \Theta : p_{ij} - p_{ji} = 0, i<j, i,j=1,\ldots,I\}} D_{\phi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\theta})) \tag{10}$$

$$D_{\phi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{MH},\phi})) = \inf_{\left\{\boldsymbol{\theta}\in\Theta:\sum_{i=1}^{I} p_{ji}-\sum_{i=1}^{I} p_{ij}=0, j=1,\ldots,I-1\right\}} D_{\phi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\theta})), \quad (11)$$

and

$$D(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi})) = \inf_{\{\boldsymbol{\theta}\in\Theta: p_{ij}p_{jI}p_{Ii}-p_{iI}p_{Ij}p_{ji}=0,\ i,j=1,\ldots,I-1\}} D_{\phi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\boldsymbol{\theta})), \quad (12)$$

respectively. They represent the natural extensions of the restricted maximum likelihood estimator given in (5)–(7), because if we consider in (9), $\phi(x) = x\log x - x + 1$ we obtain the Kullback–Leibler divergence given in (8). In this sense, the Kullback–Leibler divergence measure is a particular case of the ϕ-divergence measure and it is very natural to extend the concept of the restricted maximum likelihood estimator using the ϕ-divergence measure. The estimator obtained as a generalization of the restricted maximum likelihood estimator using the ϕ-divergence measure is called the restricted minimum ϕ-divergence estimator. More details about the restricted minimum ϕ-divergence estimator can be seen in [22]. Caussinus [10] showed that symmetry, (2), is equivalent to quasi-symmetry, (4), and marginal homogeneity, (3), simultaneously holding; thus we have

$$\text{Quasi-Symmetry} + \text{Marginal homogeneity} = \text{Symmetry}. \quad (13)$$

Thus for conditional quasi-symmetry, testing marginal homogeneity is equivalent to testing symmetry. In this work we present two new families of test statistics, based on ϕ-divergences, to define two conditional tests for marginal homogeneity taking into account relation (13). In Section 2 we present the two new families of test statistics and we obtain the asymptotic distribution.

2. Phi-divergence test statistics for testing marginal homogeneity

Menéndez et al. [21] obtained the following asymptotic expressions for the estimators $\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}$ of $\boldsymbol{\theta}_0$. For $\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}$,

$$\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi} = \boldsymbol{\theta}_0 + H_{\mathrm{QS}}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{A}(\boldsymbol{\theta}_0)^{\mathrm{T}}\mathrm{diag}\,(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\,(\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}_0)) + o_p(n^{-1/2}) \quad (14)$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0} = \mathrm{diag}(\boldsymbol{\theta}_0) - \boldsymbol{\theta}_0\boldsymbol{\theta}_0^{\mathrm{T}}, \boldsymbol{A}(\boldsymbol{\theta}_0) = \mathrm{diag}(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})(\frac{\partial \boldsymbol{p}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}})_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$,

$$\boldsymbol{H}_{\mathrm{QS}}(\boldsymbol{\theta}_0) = \boldsymbol{I}_{(I^2-1)\times(I^2-1)} - \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{B}_{\mathrm{QS}}(\boldsymbol{\theta}_0)^{\mathrm{T}}(\boldsymbol{B}_{\mathrm{QS}}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{B}_{\mathrm{QS}}(\boldsymbol{\theta}_0)^{\mathrm{T}})^{-1}\boldsymbol{B}_{\mathrm{QS}}(\boldsymbol{\theta}_0),$$

$\boldsymbol{B}_{\mathrm{QS}}(\boldsymbol{\theta}_0) = \left(\frac{\partial h_{ij}^{\mathrm{QS}}(\boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}}\right)_{(I-1)(I-2)/2\times(I^2-1)}$ and $h_{ij}^{\mathrm{QS}}(\boldsymbol{\theta}) = p_{ij}p_{jI}p_{Ii} - p_{iI}p_{Ij}p_{ji}, i, j = 1, \ldots, I-1$. For $\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}$,

$$\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi} = \boldsymbol{\theta}_0 + \boldsymbol{H}_S(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{A}\,(\boldsymbol{\theta}_0)^{\mathrm{T}}\,\mathrm{diag}\,(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\,(\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}_0)) + o_p(n^{-1/2}) \quad (15)$$

where

$$\boldsymbol{H}_S(\boldsymbol{\theta}_0) = \boldsymbol{I}_{(I^2-1)\times(I^2-1)} - \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{B}_S(\boldsymbol{\theta}_0)^{\mathrm{T}}(\boldsymbol{B}_S(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{B}_S(\boldsymbol{\theta}_0)^{\mathrm{T}})^{-1}\boldsymbol{B}_S(\boldsymbol{\theta}_0),$$

$\boldsymbol{B}_S(\boldsymbol{\theta}_0) = \left(\frac{\partial h_{ij}^{\mathrm{S}}(\boldsymbol{\theta}_0)}{\partial\theta_{ij}}\right)_{\frac{I(I-1)}{2}\times(I^2-1)}$ and $h_{ij}^{\mathrm{S}}(\boldsymbol{\theta}) = p_{ij} - p_{ji}, i, j = 1, \ldots, I$.

A similar asymptotic decomposition can be obtained for $\hat{\boldsymbol{\theta}}^{\mathrm{MH},\phi}$. We do not present it because it is not necessary in our study, but it is possible to find it in [21]. It is important to observe that the asymptotic decomposition of the estimators $\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}$ (the same happens for $\hat{\boldsymbol{\theta}}^{\mathrm{MH},\phi}$) is independent of the function ϕ considered. Then all of them have the same asymptotic properties and, of course, the same ones as the corresponding maximum likelihood estimators $\hat{\boldsymbol{\theta}}^{\mathrm{S}}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{QS}}$ because they are obtained from $\phi(x) = x\log x - x + 1$.

On the basis of (13) it is possible to test conditional marginal homogeneity by comparing the model under the assumption of quasi-symmetry and the model under the assumption of symmetry. We will consider the two following families of ϕ-divergence test statistics:

$$W^{\mathrm{MH}}_{\varphi,\phi} = \frac{2n}{\varphi''(1)} \left(D_{\varphi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi})) - D_{\varphi}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi})) \right) \tag{16}$$

and

$$S^{\mathrm{MH}}_{\varphi,\phi} = \frac{2n}{\varphi''(1)} D_{\varphi} \left(\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}), \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}) \right). \tag{17}$$

The family $W^{\mathrm{MH}}_{\varphi,\phi}$ is a natural extension of the likelihood ratio test for this problem because

$$LR = 2n \left(D_{\mathrm{Kullback}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S}})) - D_{\mathrm{Kullback}}(\hat{\boldsymbol{p}}, \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS}})) \right). \tag{18}$$

The second family, $S^{\mathrm{MH}}_{\varphi,\phi}$, is based on the following idea:

$$LR = 2n D_{\mathrm{Kullback}}(\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS}}), \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S}})) + o_p(1). \tag{19}$$

The expression given in (16) is a natural extension of the expression given in (18) and the expression given in (17) is a natural extension of the expression given in (19). It is also interesting to observe that the classical chi-squared test statistic can be obtained from (17) with $\varphi(x) = \frac{1}{2}(x-1)^2$ and $\phi(x) = x \log x - x + 1$.

In the following theorem we present the asymptotic distribution.

Theorem 1. *For testing hypotheses,*

H_0 : *Symmetry versus* H_1 : *Quasi-Symmetry*,

the asymptotic null distribution of the ϕ-divergence test statistics $W^{\mathrm{MH}}_{\varphi,\phi}$ and $S^{\mathrm{MH}}_{\varphi,\phi}$ given in (16) and (17) respectively is chi squared with $I-1$ degrees of freedom.

Proof. Firstly, we shall obtain the asymptotic distribution of the ϕ-divergence test statistic $S^{\mathrm{MH}}_{\varphi,\phi}$.

The second-order Taylor expansion of $D_{\varphi}(\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}), \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}))$ around $(\boldsymbol{p}(\boldsymbol{\theta}_0), \boldsymbol{p}(\boldsymbol{\theta}_0))$ is given by

$$\frac{2n}{\varphi''(1)} D_{\varphi}(\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}), \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi})) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} + o_p(1)$$

where $\boldsymbol{X}$ is a random vector defined by

$$\boldsymbol{X} = \sqrt{n}\,\mathrm{diag}\,(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\,(\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}) - \boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi})).$$

Then the ϕ-divergence test statistic $S^{\mathrm{MH}}_{\varphi,\phi}$ and the quadratic form $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ have the same asymptotic distribution.

The first-order Taylor expansions of $\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi})$ and $\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi})$ at $\boldsymbol{\theta}_0$ are given by

$$\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}) - \boldsymbol{p}(\boldsymbol{\theta}_0) = \frac{\partial \boldsymbol{p}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} (\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi} - \boldsymbol{\theta}_0) + o_p(\|\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi} - \boldsymbol{\theta}_0\|)$$

and

$$\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}) - \boldsymbol{p}(\boldsymbol{\theta}_0) = \frac{\partial \boldsymbol{p}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} (\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi} - \boldsymbol{\theta}_0) + o_p(\|\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi} - \boldsymbol{\theta}_0\|).$$

But, taking in account (14) and (15), we have

$$\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{QS},\phi}) - \boldsymbol{p}(\boldsymbol{\theta}_0) = \frac{\partial \boldsymbol{p}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \boldsymbol{H}_{\mathrm{QS}}(\boldsymbol{\theta}_0) \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0} \boldsymbol{A}(\boldsymbol{\theta}_0)^{\mathrm{T}} \mathrm{diag}\,(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\,(\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}_0)) + o_p(n^{-1/2}),$$

and

$$\boldsymbol{p}(\hat{\boldsymbol{\theta}}^{\mathrm{S},\phi}) - \boldsymbol{p}(\boldsymbol{\theta}_0) = \frac{\partial \boldsymbol{p}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \boldsymbol{H}_{S}(\boldsymbol{\theta}_0) \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0} \boldsymbol{A}(\boldsymbol{\theta}_0)^{\mathrm{T}} \mathrm{diag}\,(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\,(\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}_0)) + o_p(n^{-1/2}).$$

Hence,

$$\boldsymbol{X} = \sqrt{n}(\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) - \boldsymbol{L}_S(\boldsymbol{\theta}_0))\mathrm{diag}\,(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\,(\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}_0)) + o_p(1)$$

where, $\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) = \boldsymbol{A}(\boldsymbol{\theta}_0)\boldsymbol{H}_{\mathrm{QS}}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{A}(\boldsymbol{\theta}_0)^{\mathrm{T}}$ and $\boldsymbol{L}_S(\boldsymbol{\theta}_0) = \boldsymbol{A}(\boldsymbol{\theta}_0)\boldsymbol{H}_S(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\boldsymbol{A}(\boldsymbol{\theta}_0)^{\mathrm{T}}$.

Therefore, $\boldsymbol{X} \xrightarrow[n\to\infty]{L} N(\boldsymbol{0}, \boldsymbol{\Sigma}_1)$ where

$$\boldsymbol{\Sigma}_1 = (\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) - \boldsymbol{L}_S(\boldsymbol{\theta}_0))\mathrm{diag}(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathrm{diag}(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})(\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) - \boldsymbol{L}_S(\boldsymbol{\theta}_0))^{\mathrm{T}}.$$

But, $\mathrm{diag}(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2})\Sigma_{\boldsymbol{\theta}_0}\mathrm{diag}(\boldsymbol{p}(\boldsymbol{\theta}_0)^{-1/2}) = \boldsymbol{I} - \sqrt{\boldsymbol{p}(\boldsymbol{\theta}_0)}\sqrt{\boldsymbol{p}(\boldsymbol{\theta}_0)}^{\mathrm{T}}$ and $\sqrt{p(\boldsymbol{\theta}_0)}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{\theta}_0) = 0$, and thus

$$\boldsymbol{\Sigma}_1 = (\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) - \boldsymbol{L}_S(\boldsymbol{\theta}_0))(\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) - \boldsymbol{L}_S(\boldsymbol{\theta}_0))^{\mathrm{T}}.$$

It is not difficult to establish that $\boldsymbol{\Sigma}_1 = (\boldsymbol{L}_{\mathrm{QS}}(\boldsymbol{\theta}_0) - \boldsymbol{L}_S(\boldsymbol{\theta}_0))$ and this matrix is idempotent and its trace is $I - 1$. Therefore, the asymptotic distribution of $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ is chi squared with $I - 1$ degrees of freedom.
In a similar way we can obtain the asymptotic distribution of the statistic $W_{\varphi,\phi}^{\mathrm{MH}}$. □

Remark 2. If we use the ϕ-divergence test statistics $W_{\varphi,\phi}^{\mathrm{MH}}$ ($S_{\varphi,\phi}^{\mathrm{MH}}$) for testing the conditional marginal homogeneity we must reject the null hypothesis, i.e., the hypothesis of marginal homogeneity if $W_{\varphi,\phi}^{\mathrm{MH}}$ ($S_{\varphi,\phi}^{\mathrm{MH}}$) is too large. When $W_{\varphi,\phi}^{\mathrm{MH}} > c_1$ ($S_{\varphi,\phi}^{\mathrm{MH}} > c_2$) we must reject the null hypothesis of marginal homogeneity, where c_1 (c_2) is specified so that the size of the test is α:

$$\Pr(W_{\varphi,\phi}^{\mathrm{MH}} \geq c_1\ (S_{\varphi,\phi,h}^{\mathrm{MH}} \geq c_2)/H_0) = \alpha; \quad \alpha \in (0, 1).$$

On the basis of Theorem 1, the values c_1 (c_2) could be chosen as the $(1 - \alpha)$-th quantile of a chi-squared distribution with $I - 1$ degrees of freedom: $c_1\ (c_2) = \chi^2_{I-1,1-\alpha}$, where $\Pr(\chi^2_f \geq \chi^2_{f,p}) = p$.

For these tests to be valid, the quasi-symmetry model must hold true. In cases when the quasi-symmetry model is not true then the unconditional test for marginal homogeneity should be used. For more details about unconditional tests for marginal homogeneity based on ϕ-divergence test statistics see [21].

Acknowledgements

This work was partially supported by Grants DGES PB2003-892, UCM 2005-910707 and HG2004-0012 (Bilateral agreement between the Greek Ministry for Development (General Secretariat for Research and Technology) and the Spanish Ministry of Education and Science (2004–2006))

References

[1] A. Agresti, Testing marginal homogeneity for ordinal categorical variables, Biometrics 39 (1983) 505–511.
[2] A. Agresti, Categorical Data Analysis, Second edn, John Wiley and Sons, New York, 2002.
[3] S.M. Ali, S.D. Silvey, A general class of coefficients of divergence of one distribution from another, J. Roy. Statist. Soc., Ser. B 28 (1966) 131–142.
[4] E.B. Andersen, Introduction to the Statistical Analysis of Categorical Data, Springer, 1998.
[5] V.P. Bhapkar, A note on the equivalence of two criteria for hypotheses in categorical data, J. Amer. Statist. Assoc. 61 (1966) 228–235.
[6] V.P. Bhapkar, On tests of marginal symmetry and quasi-symmetry in two and three-dimensional contingency tables, Biometrics 35 (1979) 417–426.
[7] V.P. Bhapkar, J.N. Darroch, Marginal symmetry and quasi symmetry of general order, J. Multivariate Anal. 34 (1990) 173–184.
[8] M.M. Bishop, S.E. Fienberg, P.W. Holland, Discrete Multivariate Analysis: Theory and Practice, Cambridge, Mass, MIT Press, 1975.
[9] A. Bowker, A test for symmetry in contingency tables, J. Amer. Statist. Assoc. 43 (1948) 572–574.
[10] H. Caussinus, Contribution à l'analyse statistique des tableaux de correlation, Ann. Fac. Sci. Univ. Toulouse 29 (1965) 77–182.
[11] I. Csiszàr, Eine Informationstheorestiche Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, Publ. Math. Inst. Hungarian Acad. Sci. Ser. A 8 (1963) 84–108.
[12] J.N. Darroch, The Mantel–Haenszel test and tests of marginal symmetry; Fixed effects and mixed models for a categorical response, Int. Statist. Rev. 49 (1981) 285–307.
[13] J.N. Darroch, Quasi-symmetry, in: Encyclopedia of Statistical Sciences, vol. 7, John Wiley, New York, 1986, pp. 469–473.
[14] J.N. Darroch, P. McCullagh, Category distinguishability and observer agreement, Aust. J. Stat. 28 (1986) 371–388.
[15] C.T. Ireland, H.H. Ku, G.G. Koch, Symmetry and marginal homogeneity of an $r \times r$ contingency table, J. Amer. Statist. Assoc. 64 (1969) 1323–1341.

[16] S. Kullback, Information Theory and Statistics, Wiley, New York, 1959.
[17] S. Kullback, Marginal homogeneity of multidimensional contingency tables, Ann. Math. Stat. 42 (1971) 594–606.
[18] G.B. Matthews, N.A.S. Crowther, A maximum likelihood estimation procedure when modelling categorical data in terms of cross-product ratios, South African Statist. J. 31 (1997) 161–184.
[19] P. McCullagh, Some applications of quasi-symmetry, Biometrika 69 (1982) 303–308.
[20] M.L. Menéndez, J.A. Pardo, L. Pardo, Tests based on ϕ-divergences for bivariate symmetry, Metrika 53 (2001) 15–29.
[21] M.L. Menéndez, J.A. Pardo, L. Pardo, K. Zografos, On tests of symmetry, marginal homogeneity and quasi-symmetry in two contingency tables based on minimum ϕ-divergence estimator with constraints, J. Stat. Comput. Simul. 5 (7) (2005) 555–580.
[22] J.A. Pardo, L. Pardo, K. Zografos, Minimum ϕ-divergence estimators with constraints in multinomial populations, J. Statist. Plann. Inference 104 (2002) 221–237.
[23] L. Pardo, Statistical inference based on divergence measures, in: Statistics: Textbooks and Monographs, Chapman & Hall/CRC, New York, 2006.
[24] D. Quade, I.A. Salama, A note on minimum chi-square statistics in contingency tables, Biometrics 31 (1975) 953–956.
[25] J.S. Simonoff, Analyzing Categorical Data, Springer-Verlag, New York, 2003.