Conditional tests of marginal homogeneity based on $\phi$-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 $\phi$-divergence measures are introduced for testing conditional marginal homogeneity assuming that quasi-symmetry holds.

Keywords: Symmetry; Quasi-symmetry; Marginal homogeneity; $\phi$-Divergence test statistics

1. Introduction

Let $\Omega$ 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 $\Omega$ a random sample of size $n$ and define $N_{ij} = \sum_{l=1}^{n} I_{\{x_l = x_i, y_l = y_j\}}$, $i, j = 1, \ldots, I$. 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

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

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

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

∗ 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).

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Symmetry; Quasi-symmetry; Marginal homogeneity; $\phi$-Divergence test statistics

doi:10.1016/j.aml.2006.02.015
\[ H_0 : \sum_{i=1}^{I} p_{ji} = \sum_{i=1}^{I} p_{ij}, \quad j = 1, \ldots, I - 1 \]  
(3)
and
\[ H_0 : p_{ij} p_{ji} - p_{i1} p_{1j} = 0, \quad i, j = 1, \ldots, I - 1, \]  
(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 \( \phi \)-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 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, \( \hat{\theta}^S \) (Symmetry), \( \hat{\theta}^{MH} \) (Marginal Homogeneity) and \( \hat{\theta}^{QS} \) (Quasi-symmetry) are given by
\[
D_{\text{Kull}}(\hat{\theta}^S, p(\hat{\theta}^S)) = \inf_{\theta \in \Theta : p_{ji} - p_{ij} = 0, i < j, i, j = 1, \ldots, I} D_{\text{Kull}}(\hat{\theta}, p(\theta))
\]  
(5)
\[
D_{\text{Kull}}(\hat{\theta}^{MH}, p(\hat{\theta}^{MH})) = \inf_{\theta} D_{\text{Kull}}(\hat{\theta}, p(\theta))
\]  
(6)
\[
D_{\text{Kull}}(\hat{\theta}^{QS}, p(\hat{\theta}^{QS})) = \inf_{\theta \in \Theta : p_{ji} p_{ji} - p_{i1} p_{1j} = 0, i, j = 1, \ldots, I - 1} D_{\text{Kull}}(\hat{\theta}, p(\theta)),
\]  
(7)
where \( D_{\text{Kull}}(pq) \) is the Kullback–Leibler measure of divergence, see [16], between the probability vectors \( p = (p_{11}, \ldots, p_{II})^T \) and \( q = (q_{11}, \ldots, q_{II})^T \), defined by
\[
D_{\text{Kull}}(p, q) = \sum_{i=1}^{I} \sum_{j=1}^{I} p_{ij} \log \frac{p_{ij}}{q_{ij}}.
\]  
(8)

In [21] the three problems were studied using the restricted minimum \( \phi \)-divergence estimator. This estimator is based on the \( \phi \)-divergence measure defined independently by [11] and [3],
\[
D_{\phi}(p, q) = \sum_{i=1}^{I} \sum_{j=1}^{I} q_{ij} \phi \left( \frac{p_{ij}}{q_{ij}} \right); \quad \phi \in \Phi^*
\]  
(9)
where \( \Phi^* \) 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 \( x = 0, \phi(0) = 0 \) and \( \phi(p/0) = \lim_{u \to -\infty} \phi(u)/u \). For more details about \( \phi \)-divergence measures, see [23].

The restricted minimum \( \phi \)-divergence estimators for the problems considered in (2)–(4) could be obtained as the values \( \hat{\theta}^{S,\phi} \) (Symmetry), \( \hat{\theta}^{MH,\phi} \) (Marginal Homogeneity), and \( \hat{\theta}^{QS,\phi} \) (Quasi-symmetry) verifying
\[
D_{\phi}(\hat{\theta}, p(\hat{\theta}^{S,\phi})) = \inf_{\theta \in \Theta : p_{ji} - p_{ij} = 0, i < j, i, j = 1, \ldots, I} D_{\phi}(\hat{\theta}, p(\theta))
\]  
(10)
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

sence, the Kullback–Leibler divergence measure is a particular case of the \( \phi \)-divergence measure and it is very natural to extend the concept of the restricted maximum likelihood estimator using the \( \phi \)-divergence measure. The estimator obtained as a generalization of the restricted maximum likelihood estimator using the \( \phi \)-divergence measure is called the restricted minimum \( \phi \)-divergence estimator. More details about the restricted minimum \( \phi \)-divergence estimator can be seen in [22]. Casassus [10] showed that symmetry, (2), is equivalent to quasi-symmetry, (4), and marginal homogeneity, (3), simultaneously holding; thus we have

\[
\text{Quasi-Symmetry + Marginal homogeneity = Symmetry.}
\]

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 \( \phi \)-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 \( \hat{\theta}^{S,\phi} \) and \( \hat{\theta}^{QS,\phi} \) of \( \theta_0 \). For \( \hat{\theta}^{QS,\phi} \),

\[
\hat{\theta}^{QS,\phi} = \theta_0 + H_{QS}(\theta_0) \Sigma_{\theta_0} A(\theta_0)^T \text{ diag } (p(\theta_0)^{-1/2}) (\hat{\theta} - p(\theta_0)) + o_p(n^{-1/2})
\]

where \( \Sigma_{\theta_0} = \text{ diag } (\theta_0 - \theta_0^T) \), \( A(\theta_0) = \text{ diag } (p(\theta_0)^{-1/2}) (\frac{\partial p(\theta)}{\partial \theta}) \theta_0 - \theta_0 \),

\[
H_{QS}(\theta_0) = I_{(I^2-1) \times (I^2-1)} - \Sigma_{\theta_0} B_{QS}(\theta_0)^T (B_{QS}(\theta_0) \Sigma_{\theta_0} B_{QS}(\theta_0)^T)^{-1} B_{QS}(\theta_0),
\]

\[
B_{QS}(\theta_0) = \left( \frac{\partial h_{ij}^{QS}(\theta_0)}{\partial \theta_j} \right)_{(I-1)(I-2)/2 \times (I^2-1)} \quad \text{and} \quad h_{ij}^{QS}(\theta) = p_{ij} p_{ji} - p_{ij} p_{ji}, i, j = 1, \ldots, I - 1.
\]

For \( \hat{\theta}^{S,\phi} \),

\[
\hat{\theta}^{S,\phi} = \theta_0 + H_S(\theta_0) \Sigma_{\theta_0} A(\theta_0)^T \text{ diag } (p(\theta_0)^{-1/2}) (\hat{\theta} - p(\theta_0)) + o_p(n^{-1/2})
\]

where

\[
H_S(\theta_0) = I_{(I^2-1) \times (I^2-1)} - \Sigma_{\theta_0} B_S(\theta_0)^T (B_S(\theta_0) \Sigma_{\theta_0} B_S(\theta_0)^T)^{-1} B_S(\theta_0),
\]

\[
B_S(\theta_0) = \left( \frac{\partial h_{ij}^S(\theta_0)}{\partial \theta_j} \right)_{(I-1)(I-2)/2 \times (I^2-1)} \quad \text{and} \quad h_{ij}^S(\theta) = p_{ij} - p_{ji}, i, j = 1, \ldots, I.
\]

A similar asymptotic decomposition can be obtained for \( \hat{\theta}^{\text{MH,}\phi} \). 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 \( \hat{\theta}^{S,\phi} \) and \( \hat{\theta}^{QS,\phi} \) (the same happens for \( \hat{\theta}^{\text{MH,}\phi} \)) is independent of the function \( \phi \) considered. Then all of them have the same asymptotic properties and, of course, the same ones as the corresponding maximum likelihood estimators \( \hat{\theta}^{S} \) and \( \hat{\theta}^{QS} \) 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 $\phi$-divergence test statistics:

$$W_{\phi,\phi}^{\text{MH}} = \frac{2n}{\phi''(1)} \left( D_{\phi}(\hat{\theta}, p(\hat{\theta}^{S,\phi})) - D_{\phi}(\hat{\theta}, p(\hat{\theta}^{QS,\phi})) \right)$$

(16)

and

$$S_{\phi,\phi}^{\text{MH}} = \frac{2n}{\phi''(1)} D_{\phi} \left( p(\hat{\theta}^{QS,\phi}), p(\hat{\theta}^{S,\phi}) \right).$$

(17)

The family $W_{\phi,\phi}^{\text{MH}}$ is a natural extension of the likelihood ratio test for this problem because

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

(18)

The second family, $S_{\phi,\phi}^{\text{MH}}$, is based on the following idea:

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

(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 $\phi(x) = \frac{1}{2} (x - 1)^2$ and $\phi(x) = x \log x - x + 1$.

In the following theorem we present the asymptotic distribution.

**Theorem 1.** For testing hypotheses, $H_0 : \text{Symmetry versus } H_1 : \text{Quasi-Symmetry},$

the asymptotic null distribution of the $\phi$-divergence test statistics $W_{\phi,\phi}^{\text{MH}}$ and $S_{\phi,\phi}^{\text{MH}}$ 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 $\phi$-divergence test statistic $S_{\phi,\phi}^{\text{MH}}$. The second-order Taylor expansion of $D_{\phi}(p(\hat{\theta}^{QS,\phi}), p(\hat{\theta}^{S,\phi}))$ around $(p(\theta_0), p(\theta_0))$ is given by

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

where $X$ is a random vector defined by

$$X = \sqrt{n} \text{diag} (p(\theta_0)^{-1/2}) (p(\hat{\theta}^{QS,\phi}) - p(\hat{\theta}^{S,\phi})).$$

Then the $\phi$-divergence test statistic $S_{\phi,\phi}^{\text{MH}}$, and the quadratic form $X^T X$ have the same asymptotic distribution.

The first-order Taylor expansions of $p(\hat{\theta}^{QS,\phi})$ and $p(\hat{\theta}^{S,\phi})$ at $\theta_0$ are given by

$$p(\hat{\theta}^{QS,\phi}) - p(\theta_0) = \frac{\partial p(\theta_0)}{\partial \theta} (\theta^{QS,\phi} - \theta_0) + o_p(\|\theta^{QS,\phi} - \theta_0\|)$$

and

$$p(\hat{\theta}^{S,\phi}) - p(\theta_0) = \frac{\partial p(\theta_0)}{\partial \theta} (\theta^{S,\phi} - \theta_0) + o_p(\|\theta^{S,\phi} - \theta_0\|).$$

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

$$p(\hat{\theta}^{QS,\phi}) - p(\theta_0) = \frac{\partial p(\theta_0)}{\partial \theta} H_{QS}(\theta_0) \Sigma_{\theta_0} A(\theta_0)^T \text{diag} (p(\theta_0)^{-1/2}) (\hat{\theta} - p(\theta_0)) + o_p(n^{-1/2}),$$

and

$$p(\hat{\theta}^{S,\phi}) - p(\theta_0) = \frac{\partial p(\theta_0)}{\partial \theta} H_{S}(\theta_0) \Sigma_{\theta_0} A(\theta_0)^T \text{diag} (p(\theta_0)^{-1/2}) (\hat{\theta} - p(\theta_0)) + o_p(n^{-1/2}).$$
Hence, \[ X = \sqrt{n}(L_{QS}(\theta_0) - L_S(\theta_0)) \text{diag}(p(\theta_0)^{-1/2})(\hat{p} - p(\theta_0)) + o_p(1) \]
where, \( L_{QS}(\theta_0) = A(\theta_0)H_{QS}(\theta_0)\Sigma_{\theta_0}A(\theta_0)^T \) and \( L_S(\theta_0) = A(\theta_0)H_S(\theta_0)\Sigma_{\theta_0}A(\theta_0)^T \).

Therefore, \( X \overset{L}{\rightarrow} N(0, \Sigma_1) \) where
\[ \Sigma_1 = (L_{QS}(\theta_0) - L_S(\theta_0)) \text{diag}(p(\theta_0)^{-1/2})\Sigma_{\theta_0}\text{diag}(p(\theta_0)^{-1/2})(L_{QS}(\theta_0) - L_S(\theta_0))^T. \]

But, \( \text{diag}(p(\theta_0)^{-1/2})\Sigma_{\theta_0}\text{diag}(p(\theta_0)^{-1/2}) = I - \sqrt{p(\theta_0)^T/p(\theta_0)} \text{ and } \sqrt{p(\theta_0)^T}A(\theta_0) = 0, \)
and thus
\[ \Sigma_1 = (L_{QS}(\theta_0) - L_S(\theta_0))(L_{QS}(\theta_0) - L_S(\theta_0))^T. \]

It is not difficult to establish that \( \Sigma_1 = (L_{QS}(\theta_0) - L_S(\theta_0)) \) and this matrix is idempotent and its trace is \( I - 1 \).

Therefore, the asymptotic distribution of \( X^T X \) is chi squared with \( I - 1 \) degrees of freedom.

In a similar way we can obtain the asymptotic distribution of the statistic \( W_{\psi,\phi}^{MH} \). □

**Remark 2.** If we use the \( \psi \)-divergence test statistics \( W_{\psi,\phi}^{MH} (S_{\psi,\phi}^{MH}) \) for testing the conditional marginal homogeneity we must reject the null hypothesis, i.e., the hypothesis of marginal homogeneity if \( W_{\psi,\phi}^{MH} (S_{\psi,\phi}^{MH}) \) is too large. When \( W_{\psi,\phi}^{MH} > c_1 (S_{\psi,\phi}^{MH} > c_2) \) we must reject the null hypothesis of marginal homogeneity, where \( c_1 (c_2) \) is specified so that the size of the test is \( \alpha \):
\[ \text{Pr}(W_{\psi,\phi}^{MH} \geq c_1 (S_{\psi,\phi}^{MH} \geq c_2) / H_0) = \alpha; \quad \alpha \in (0, 1). \]

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: \( c_1 (c_2) = \chi_{I-1,1-\alpha}^2, \) where \( \text{Pr}(\chi_f^2 \geq \chi_{I,f,p}^2) = p. \)

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 \( \psi \)-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**