Journal of Computational and Applied Mathematics 235 (2011) 1160–1179

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

On the comparison of the pre-test and shrinkage phi-divergence test estimators for the symmetry model of categorical data

Leandro Pardo [a,*], Nirian Martín [b]

[a] Department of Statistics and O.R., Complutense University of Madrid, 28040 Madrid, Spain
[b] Department of Statistics, Carlos III University of Madrid, 28903 Getafe (Madrid), Spain

ARTICLE INFO

Article history:
Received 6 September 2009
Received in revised form 25 February 2010

MSC:
62B10
62H15

Keywords:
Minimum phi-divergence estimator
Phi-divergence statistics
Preliminary test estimator
Symmetry model

ABSTRACT

The estimation problem of the parameters in a symmetry model for categorical data has been considered for many authors in the statistical literature (for example, Bowker (1948) [1], Ireland et al. (1969) [2], Quade and Salama (1975) [3], Cressie and Read (1988) [4], Menéndez et al. (2005) [5]) without using uncertain prior information. It is well known that many new and interesting estimators, using uncertain prior information, have been studied by a host of researchers in different statistical models, and many papers have been published on this topic (see Saleh (2006) [9] and references therein). In this paper, we consider the symmetry model of categorical data and we study, for the first time, some new estimators when non-sample information about the symmetry of the probabilities is considered. The decision to use a "restricted" estimator or an "unrestricted" estimator is based on the outcome of a preliminary test, and then a shrinkage technique is used. It is interesting to note that we present a unified study in the sense that we consider not only the maximum likelihood estimator and likelihood ratio test or chi-square test statistic but we consider minimum phi-divergence estimators and phi-divergence test statistics. Families of minimum phi-divergence estimators and phi-divergence test statistics are wide classes of estimators and test statistics that contain as a particular case the maximum likelihood estimator, likelihood ratio test and chi-square test statistic. In an asymptotic set-up, the biases and the risk under the squared loss function for the proposed estimators are derived and compared. A numerical example clarifies the content of the paper.

1. Introduction

Let X and Y denote two ordinal response variables, X and Y having I levels. When we classify subjects on both variables, there are I^2 possible combinations of classifications. The responses (X, Y) of a subject randomly chosen from some population have a probability distribution. Let $p_{ij} = P(X = i, Y = j)$, with $p_{ij} > 0$, $i, j = 1, \ldots, I$. We display this distribution in a rectangular table having I rows for the categories of X and I columns for the categories of Y. We denote by $\boldsymbol{p} = (p_{11}, \ldots, p_{II})^T$ the $I^2 - 1$ unknown parameters in the model. In order to get the "unrestricted" estimator of $\boldsymbol{p}$ we consider a random sample of size n, $(X_1, Y_1), \ldots, (X_n, Y_n)$ from (X, Y), and we denote by

$$N_{ij} = \sum_{l=1}^{n} I_{\{i,j\}}(X_l, Y_l) \tag{1}$$

and n_{ij} a particular result of N_{ij}, i.e., n_{ij} represents the observed frequency in the (i, j)th cell for $(i, j) \in I \times I$ with $\sum_{i=1}^{I} \sum_{j=1}^{I} n_{ij} = n$. We shall denote by $\widehat{\boldsymbol{p}} = (n_{11}/n, \ldots, n_{II}/n)^T$ the "unrestricted" estimator of $\boldsymbol{p}$.

* Corresponding author. Tel.: +34 913944425.
E-mail address: lpardo@mat.ucm.es (L. Pardo).

0377-0427/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2010.07.026

Also assume that uncertain non-sample prior information on the value of $\boldsymbol{p} = (p_{11}, \dots, p_{II})^T$ is available, either from previous studies or from practical experience of the researchers or experts. Let the non-sample prior information be expressed in the form of the null hypothesis

$$H_0 : p_{ij} = p_{ji} \quad i, j \in I \times I \tag{2}$$

which could be true, but not necessarily. We denote by $\widehat{\boldsymbol{p}}_R$ a "restricted" estimator of $\boldsymbol{p} = (p_{11}, \dots, p_{II})^T$ obtained under assumption (2). Hypothesis (2) represents the symmetry model. To investigate if there are symmetric patterns in the data is equivalent to studying if the cell probabilities on one side of the main diagonal are a mirror image of those on the other side. Some well-known "restricted" estimators (estimators under the hypothesis of symmetry), in a contingency table, have been given in [1–5], amongst others. Later we shall give the expression of the some "restricted" estimators of $\boldsymbol{p} = (p_{11}, \dots, p_{II})^T$. These estimators are based exclusively on the sample data. They do not use any other prior knowledge in their definitions. We can consider no-sample information in order to improve the quality of the estimators and we can expect a better estimation. Prior information is available in the form given in (2). But it is not certain if (2) is true. To remove this uncertainty of the prior information in the model, it is natural to perform a preliminary test on the validity of the uncertain prior information.

Based on a "unrestricted" estimator and in a "restricted" estimator of $\boldsymbol{p} = (p_{11}, \dots, p_{II})^T$, we can obtain some test statistics for testing

$$H_0 : p_{ij} = p_{ji} \quad i, j \in I \times I \text{ versus } H_1 : p_{ij} \neq p_{ji}, \text{ for at least one } (i, j) \text{ pair.} \tag{3}$$

The most common test statistics for testing (3) are based on "distances" between the "unrestricted" and one "restricted" estimator of $\boldsymbol{p} = (p_{11}, \dots, p_{II})^T$,

$$d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right). \tag{4}$$

In this paper, the test statistics $d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)$, under consideration, has asymptotically a chi-square distribution with $m = I(I-1)/2$ degrees of freedom. Some examples of "distances" of the type given in (4) are Pearson's statistic, [1],

$$X^2 = \sum_{\substack{i,j \\ i<j}} \frac{\left(n_{ij} - n_{ji}\right)^2}{n_{ij} + n_{ji}} \tag{5}$$

which for large n has a chi-square distribution with $m = I(I-1)/2$ degrees of freedom, and the likelihood ratio test statistic given by

$$G^2 = 2 \sum_{\substack{i,j \\ i<j}} n_{ij} \log \frac{2n_{ij}}{n_{ji} + n_{ij}}. \tag{6}$$

Its asymptotic distribution coincides with the asymptotic distribution of X^2. Later we will see another family of test statistics of the type (4) for testing (3).

Now we can use the sample information and the non-sample information as well as the test statistic (4) to define an estimator of $\boldsymbol{p}$. Based on $\widehat{\boldsymbol{p}}$, $\widehat{\boldsymbol{p}}_R$ and $d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)$, the preliminary test estimator, $\widehat{\boldsymbol{p}}^{\text{pte}}$, is given by

$$\widehat{\boldsymbol{p}}^{\text{pte}} = \widehat{\boldsymbol{p}}_R I_{\left(0, \chi^2_{m,\alpha}\right)} \left(d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)\right) + \widehat{\boldsymbol{p}} I_{\left[\chi^2_{m,\alpha}, \infty\right)} \left(d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)\right). \tag{7}$$

By $I_A(x)$ we are denoting the indicator function taking the value 1 if $x \in A$ and 0 if $x \notin A$ and $m = I(I-1)/2$.

We can observe that "$\widehat{\boldsymbol{p}}^{\text{pte}}$" depends on the preselected level of significance of the test. To overcome this problem, we can consider the shrinkage or James–Stein estimator, $\widehat{\boldsymbol{p}}^{s}$, of $\boldsymbol{p}$, as follows:

$$\widehat{\boldsymbol{p}}^{s} = \widehat{\boldsymbol{p}}_R + \left(1 - (m-2)\, d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)^{-1}\right) \left(\widehat{\boldsymbol{p}} - \widehat{\boldsymbol{p}}_R\right). \tag{8}$$

The previous estimators ("restricted", "unrestricted", "$\widehat{\boldsymbol{p}}^{\text{pte}}$", "$\widehat{\boldsymbol{p}}^{s}$") as well as other very interesting estimators, which we shall study later, can be obtained from the family of estimators

$$\widehat{\boldsymbol{p}}^{h} = \widehat{\boldsymbol{p}}_R + \left(1 - h\left(d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)\right)\right) \left(\widehat{\boldsymbol{p}} - \widehat{\boldsymbol{p}}_R\right), \tag{9}$$

for an appropriate real function h. For instance, if we consider $h(x) = I_{\left(0, \chi^2_{m,\alpha}\right)}(x)$ we get (7), and for $h(x) = (m-2)x^{-1}$ we get (8).

Many papers have been published studying estimators of the type given in (9), for different statistical models, following the seminal work of Bancroft [6] and later Han and Bancroft [7]. They developed the preliminary test estimators that use uncertain non-sample prior information (not in the form of prior distributions), in addition to the sample information. Stein [8] introduced the Stein-rule (shrinkage) estimator for a multivariate normal population that dominates the usual maximum likelihood estimators under the square error loss function. In order to have a clear idea about the importance of this area as well as its application in different problems, see [9].

In this paper, we shall study the family of estimators (9) for appropriate estimators $\widehat{\boldsymbol{p}}_R$ and test statistics $d\left(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R\right)$ that we will describe in the following section. Section 3 is devoted to presenting some asymptotic distributional results. We do not know the existence of any paper studying preliminary test estimators for the symmetry model in categorical data.

2. Phi-divergence estimators

We consider the set

$$\Theta = \left\{ \boldsymbol{\theta} : \boldsymbol{\theta} = \left(p_{ij};\ 1 \le i \le I,\ 1 \le j \le I,\ (i,j) \neq (I,I)\right) \text{ with } p_{ij} > 0 \text{ and } \sum_{i=1}^{I} \sum_{\substack{j=1 \\ (i,j)\neq(I,I)}}^{I} p_{ij} < 1 \right\} \tag{10}$$

and we denote $\boldsymbol{p}(\boldsymbol{\theta}) = (p_{11}, \ldots, p_{II})^T = \boldsymbol{p}$, $p_{II} = 1 - \sum_{i=1}^{I} \sum_{\substack{j=1 \\ (i,j)\neq(I,I)}}^{I} p_{ij}$.

We define

$$B = \left\{ (a_{11}, \ldots, a_{1I},\ a_{22}, \ldots, a_{2I}, \ldots, a_{I-1I-1}, a_{I-1I})^T \in \mathbb{R}^{\frac{I(I+1)}{2}-1} : \sum_{i \le j} a_{ij} < 1,\ 0 < a_{ij},\ i,j = 1, \ldots, I \right\}.$$

Hypothesis (2) can be written as

$$H_0 : \boldsymbol{\theta} = \boldsymbol{g}(\boldsymbol{\beta}), \quad \boldsymbol{\beta} = (p_{11}, \ldots, p_{1I}, p_{22}, \ldots, p_{2I}, \ldots, p_{I-1I-1}, p_{I-1I})^T \in B, \tag{11}$$

where the function $\boldsymbol{g}$ is defined by $\boldsymbol{g} = \left(g_{ij};\ i,j = 1, \ldots, I,\ (i,j) \neq (I,I)\right)$ with

$$g_{ij}(\boldsymbol{\beta}) = \begin{cases} p_{ij} & i \le j \\ p_{ji} & i > j \end{cases}, \quad i,j = 1, \ldots, I-1,$$

and

$$\begin{aligned} g_{Ij}(\boldsymbol{\beta}) &= p_{jI}, \quad j = 1, \ldots, I-1 \\ g_{iI}(\boldsymbol{\beta}) &= p_{iI}, \quad i = 1, \ldots, I-1. \end{aligned}$$

Note that $\boldsymbol{p}\left(\boldsymbol{g}(\boldsymbol{\beta})\right) = \left(g_{ij}(\boldsymbol{\beta});\ i,j = 1, \ldots, I\right)^T$, where

$$g_{II}\left(\boldsymbol{\beta}\right) = 1 - \sum_{\substack{i,j=1 \\ (i,j)\neq(I,I)}}^{I} g_{ij}(\boldsymbol{\beta}).$$

The maximum likelihood estimator (MLE) of $\boldsymbol{\beta}$ can be defined as

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}\in B} D\left(\widehat{\boldsymbol{p}}, \boldsymbol{p}\left(\boldsymbol{g}(\boldsymbol{\beta})\right)\right) \quad \text{a.s.},$$

where $D\left(\widehat{\boldsymbol{p}}, \boldsymbol{p}\left(\boldsymbol{g}(\boldsymbol{\beta})\right)\right)$ is the Kullback–Leibler divergence measure defined by

$$D\left(\widehat{\boldsymbol{p}}, \boldsymbol{p}\left(\boldsymbol{g}(\boldsymbol{\beta})\right)\right) = \sum_{i=1}^{I} \sum_{j=1}^{I} \widehat{p}_{ij} \log \frac{\widehat{p}_{ij}}{g_{ij}(\boldsymbol{\beta})}.$$

We denote by $\widehat{\boldsymbol{\theta}} = \boldsymbol{g}(\widehat{\boldsymbol{\beta}})$ the MLE of $\boldsymbol{\theta} = \boldsymbol{g}(\boldsymbol{\beta})$ and by $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}) = \left(p_{11}(\widehat{\boldsymbol{\theta}}), \ldots, p_{II}(\widehat{\boldsymbol{\theta}})\right)^T$ the MLE of $\boldsymbol{p}(\boldsymbol{\theta})$. It is well known that $\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}) = \frac{\widehat{p}_{ij}+\widehat{p}_{ji}}{2}$, $i = 1, \ldots, I, j = 1, \ldots, I$. Using the ideas developed in [10], we can consider the minimum ϕ_2-divergence estimator ($M\phi_2E$) replacing the Kullback–Leibler divergence by a ϕ_2-divergence measure in the following way:

$$\widehat{\boldsymbol{\beta}}^{\phi_2} = \arg\min_{\boldsymbol{\beta}\in B} D_{\phi_2}\left(\widehat{\boldsymbol{p}}, \boldsymbol{p}\left(\boldsymbol{g}(\boldsymbol{\beta})\right)\right); \quad \phi_2 \in \Phi^*, \tag{12}$$

where

$$D_{\phi_2}\left(\widehat{\boldsymbol{p}}, \boldsymbol{p}\left(\boldsymbol{g}(\boldsymbol{\beta})\right)\right) = \sum_{i=1}^{I} \sum_{j=1}^{I} g_{ij}(\boldsymbol{\beta}) \phi_2\left(\frac{\widehat{p}_{ij}}{g_{ij}(\boldsymbol{\beta})}\right),$$

Φ^* is the class of all convex functions $\phi_2(x)$, $x > 0$, such that at $x = 1$, $\phi_2(1) = 0$, $\phi_2''(1) > 0$, and at $x = 0$, $0\phi_2\left(0/0\right) = 0$ and $0\phi_2\left(p/0\right) = p \lim_{u\to\infty} \phi_2(u)/u$. The ϕ_2-divergence measures were introduced simultaneously in [11,12]. For more details about ϕ-divergence measures, see [13] and the references therein. In the following, we shall assume that the functions ϕ_2 in the class Φ^* are twice continuously differentiable at $x > 0$.

We denote by $\widehat{\boldsymbol{\theta}}^{\phi_2} = \boldsymbol{g}(\widehat{\boldsymbol{\beta}}^{\phi_2})$ and by

$$\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \left(p_{11}(\widehat{\boldsymbol{\theta}}^{\phi_2}), \ldots, p_{II}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)^T \tag{13}$$

the $(M\phi_2E)$ of the probability vector that characterizes the symmetry model. Based on $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, it is possible to define a new family of statistics for testing (3) that contains as a particular case the test statistics given in (5) and (6). This family of test statistics is given by

$$T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \equiv \frac{2n}{\phi_1''(1)} D_{\phi_1}\left(\widehat{\boldsymbol{p}}, \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) = \frac{2n}{\phi_1''(1)} \sum_{i=1}^{I}\sum_{j=1}^{I} p_{ij}(\widehat{\boldsymbol{\theta}}^{\phi_2})\phi_1\left(\frac{\widehat{p}_{ij}}{p_{ij}(\widehat{\boldsymbol{\theta}}^{\phi_2})}\right). \tag{14}$$

We can observe that the family (14) involves two functions ϕ_1 and $\phi_2 \in \Phi^*$. We use the function ϕ_2 to obtain the $(M\phi_2E)$ and ϕ_1 to obtain the family of test statistics. The asymptotic distribution of $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ is chi-squared with $m = I(I-1)/2$ degrees of freedom (see Chapter 8 in [13]). Thus, for a given level of significance $\alpha \in (0, 1)$, the critical value of $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ may be approximated by $\chi^2_{m,\alpha}$, the upper $100\alpha\%$ of the chi-square distribution with m degrees of freedom. If, in (13), we consider $\phi_2(x) = x\log x - x + 1$, we get the Kullback–Leibler divergence, and therefore the corresponding $M\phi_2E$ is the MLE. If, in addition, we consider $\phi_1(x) = x\log x - x + 1$ or $\phi_1(x) = (x-1)^2/2$, we obtain the statistics given in (5) and (6), respectively.

When the hypothesis of symmetry holds, $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ has a smaller risk (with a quadratic loss) than $\widehat{\boldsymbol{p}}$. If the hypothesis of symmetry does not hold, the risk of $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ may go to $+\infty$, as the sample size n increases. For this reason, when the prior knowledge about the hypothesis of symmetry in (2) is rather uncertain, it may be desirable to use a preliminary test estimator. We shall consider in this paper the family of preliminary phi-divergence test estimators

$$\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) + \left(1 - h\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right). \tag{15}$$

This family of test statistics is similar to the family (14) with $\widehat{\boldsymbol{p}}_R \equiv \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ and $d(\widehat{\boldsymbol{p}}, \widehat{\boldsymbol{p}}_R) \equiv T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$. We can observe the following.

(i) For $h(x) = I_{\left(0,\chi^2_{m,\alpha}\right)}(x)$, we get $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \equiv \boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*preliminary phi-divergence test estimator*", with

$$\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) I_{\left(0,\chi^2_{m,\alpha}\right)}(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})) + \widehat{\boldsymbol{p}} I_{\left[\chi^2_{m,\alpha},\infty\right)}(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})). \tag{16}$$

Looking at $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, it is found that as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \infty$, $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \widehat{\boldsymbol{p}}$, while as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to 0$, $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$.

(ii) For $h(x) = 0$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \widehat{\boldsymbol{p}}$, the "*unrestricted estimator*".

(iii) For $h(x) = 1$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*restricted phi-divergence estimator*".

(iv) For $h(x) = 1 - a$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}^{\text{sre}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*shrinkage phi-divergence estimator*", with

$$\boldsymbol{p}^{\text{sre}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \equiv (1-a)\,\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) + a\widehat{\boldsymbol{p}}.$$

(v) For $h(x) = (1-a)\,I_{\left(0,\chi^2_{m,\alpha}\right)}(x)$, $a \in [0,1]$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*shrinkage preliminary phi-divergence test estimator*", with

$$\boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \widehat{\boldsymbol{p}} I_{\left[\chi^2_{m,\alpha},\infty\right)}(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})) + \left[a\widehat{\boldsymbol{p}} + (1-a)\,\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right] I_{\left(0,\chi^2_{m,\alpha}\right)}(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})).$$

(vi) For $h(x) = (m-2)\,x^{-1}$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}^s_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*phi-divergence James–Stein estimator*", with

$$\boldsymbol{p}^s_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) + \left(1 - (m-2)\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})^{-2}\right)\right)\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right).$$

We can see that $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ depends on the level of significance of the test. To overcome this difficulty, we can consider the James–Stein estimator. We can observe that $\boldsymbol{p}^s_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \widehat{\boldsymbol{p}}$ as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \infty$, but as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to 0$, $\boldsymbol{p}^s_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ gives inadmissible values. To avoid the inadmissibility of $\boldsymbol{p}^s_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, we are going to define another James–Stein type estimator in (vii).

(vii) For $h(x) = 1 - \left[1 - (m-2)\,x^{-1}\right] I_{(m-2,\infty)}(x)$, $(m > 2)$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \equiv \boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*phi-divergence positive part of the Stein-rule estimator*", with

$$\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) + \left(1 - (m-2)\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)^{-1}\right) I_{(m-2,\infty)}\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right).$$

We can observe that as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \infty$, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \widehat{\boldsymbol{p}}$, while as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to m-2$, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ and if, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}})$ then $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \to \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$.

(viii) For $h(x) = 1 - \left[1 - (m-2)\,x^{-1}\right] I_{\left[\chi^2_{m,\alpha},\infty\right)}(x)$, $(m > 2)$, $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) \equiv \boldsymbol{p}^{\text{pte}_+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, the "*modified preliminary phi-divergence test estimator*", with

$$\boldsymbol{p}^{\text{pte}_+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) + \left(1 - (m-2)\,T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) I_{\left[\chi^2_{m,\alpha},\infty\right)}\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right).$$

Preliminary phi-divergence estiators have been studied in different statistical problems, see for instance [14] and [15] and references therein.

3. Asymptotic bias and asymptotic quadratic risk

The Fisher information matrix of $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ ($\boldsymbol{\Theta}$ was defined in (10)), in the symmetry model, is the $\left(I^2-1\right) \times \left(I^2-1\right)$ matrix given by

$$\boldsymbol{I}_F^S(\boldsymbol{\theta}) = \boldsymbol{\Sigma}_{\boldsymbol{\theta}} - \boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{B}(\boldsymbol{\theta})^T\left(\boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{B}(\boldsymbol{\theta})^T\right)^{-1}\boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}},$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \operatorname{diag}(\boldsymbol{\theta}) - \boldsymbol{\theta}\boldsymbol{\theta}^T$ and $\boldsymbol{B}(\boldsymbol{\theta}) = \left(\frac{\partial h_{ij}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_{ij}}\right)_{\frac{I(I-1)}{2}\times\left(I^2-1\right)}$. The functions h_{ij} are

$$h_{ij}(\boldsymbol{\theta}) = p_{ij} - p_{ji}, \quad i<j,\ i = 1,\ldots,I-1,\ j = 1,\ldots,I.$$

For more details see Chapter 8 in [13]. It is not difficult to establish that the Fisher information matrix, $\boldsymbol{I}_F^S(\boldsymbol{\theta})$, can be written as

$$\boldsymbol{I}_F^S(\boldsymbol{\theta}) = \boldsymbol{M}_{\boldsymbol{\beta}}^T\boldsymbol{I}_F(\boldsymbol{\beta})^{-1}\boldsymbol{M}_{\boldsymbol{\beta}},$$

where $\boldsymbol{I}_F(\boldsymbol{\beta})$ is the Fisher information matrix corresponding to $\boldsymbol{\beta} \in B$.

We consider a contiguous sequence of alternative hypotheses that approaches the null hypothesis $H_0 : \boldsymbol{\theta} = \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))$, for some unknown $\boldsymbol{\beta} \in B$, at the rate $O\left(n^{-1/2}\right)$. Consider the multinomial probability vector

$$p_{n,ij} = \boldsymbol{p}_{ij}(\boldsymbol{g}(\boldsymbol{\beta})) + d_{ij}n^{-1/2}, \quad i = 1,\ldots,I,\ j = 1,\ldots,I,$$

where $\boldsymbol{d} = (d_{11},\ldots,d_{II})^T$ is a fixed $I^2 \times 1$ vector such that $\sum_{i=1}^I\sum_{j=1}^I d_{ij} = 0$; recall that n is the total count parameter of the multinomial distribution and $\boldsymbol{\beta} \in B$. As $n \to \infty$, the sequence of multinomial probabilities $\{\boldsymbol{p}_n\}_{n\in N}$ with $\boldsymbol{p}_n = \left(p_{n,ij},\ i = 1,\ldots,I,\ j = 1,\ldots,I\right)^T$, converges to a multinomial probability in H_0 at the rate of $O\left(n^{-1/2}\right)$. Let

$$H_{1,n} : \boldsymbol{p}_n = \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta})) + \boldsymbol{d}n^{-1/2}, \quad \boldsymbol{\beta} \in B \tag{17}$$

be a sequence of contiguous alternative hypotheses, here contiguous to the null hypothesis $H_0 : \boldsymbol{\theta} = \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))$, for some unknown $\boldsymbol{\beta} \in B$. We can observe that $\boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))$ with $\boldsymbol{\beta} \in B$ is given by $\boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta})) = \left(p_{ij,}i,j = 1,\ldots,I;\, p_{ij} = p_{ji}\right)^T$. We shall denote

$$\Theta_0 = \{\boldsymbol{\theta} \in \Theta : \boldsymbol{\theta} = \boldsymbol{g}(\boldsymbol{\beta}) \text{ for some } \boldsymbol{\beta} \in B\}.$$

In the next theorem, we get the asymptotic distribution of the statistics $\widehat{\boldsymbol{p}}, \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}), \widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ and $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ under the contiguous alternative hypothesis given in (17). These asymptotic distributions will be important to get the asymptotic bias and asymptotic quadratic bias, as well as the asymptotic distributional quadratic risk of the family of estimators defined in (15).

In the next theorem, and in the rest of the paper, by $\boldsymbol{I}_{a\times a}$ we denote the identity matrix of order a and by $\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}$ the diagonal matrix with elements $\boldsymbol{p}(\boldsymbol{\theta})$.

Theorem 1. *Under $H_{1,n}$, given $\boldsymbol{\theta} = \boldsymbol{g}(\boldsymbol{\beta})$, $\boldsymbol{\beta} \in B$, we have the following.*
(a) *The random vector $\boldsymbol{X}_n = \sqrt{n}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}))$ converges in law to the I^2-dimensional normal random vector, $\boldsymbol{X}$, with mean vector $\delta = \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\boldsymbol{d}$ and variance–covariance matrix $\boldsymbol{J}(\boldsymbol{\theta}) = \boldsymbol{I}_{I^2\times I^2} - \sqrt{\boldsymbol{p}(\boldsymbol{\theta})}\left(\sqrt{\boldsymbol{p}(\boldsymbol{\theta})}\right)^T$.*
(b) *The random vector $\boldsymbol{Y}_n = \sqrt{n}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$ converges in law to the I^2-dimensional normal random variable, $\boldsymbol{Y}$, with mean vector $\boldsymbol{J}^*(\boldsymbol{\theta})\,\delta$, where $\boldsymbol{J}^*(\boldsymbol{\theta}) = \boldsymbol{I}_{I^2\times I^2} - \boldsymbol{K}(\boldsymbol{\theta})$ and variance–covariance matrix $\boldsymbol{B}(\boldsymbol{\theta}) = \boldsymbol{J}(\boldsymbol{\theta}) - \boldsymbol{K}(\boldsymbol{\theta})$, with $\boldsymbol{K}(\boldsymbol{\theta}) = \boldsymbol{A}(\boldsymbol{\theta})\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{A}(\boldsymbol{\theta})^T$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})} - \boldsymbol{p}(\boldsymbol{\theta})\,\boldsymbol{p}(\boldsymbol{\theta})^T$ and*

$$\boldsymbol{A}(\boldsymbol{\theta}) = \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\boldsymbol{T},$$

where $\boldsymbol{T} = \left(\boldsymbol{I}_{I^2-1\times I^2-1}, -\mathbf{1}\right)$ and $-\mathbf{1} = \left(-1, \overset{(I^2-1)}{\ldots}, -1\right)^T$.
(c) *The random vector $\boldsymbol{Z}_n = \sqrt{n}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}(\boldsymbol{\theta})\right)$ converges in law to the I^2-dimensional normal random vector, $\boldsymbol{Z}$, with mean vector $\boldsymbol{K}(\boldsymbol{\theta})\delta$ and variance–covariance matrix $\boldsymbol{K}(\boldsymbol{\theta})$.*

(d) *The family of test statistics* $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ *can be written as*

$$T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{Y}_n^T \boldsymbol{Y}_n + o_p(1),$$

and its asymptotic distribution, under the contiguous hypothesis given in (17), is a noncentral chi-squared distribution with $m = I(I-1)/2$ *degrees of freedom and non-centrality parameter*

$$\lambda = \boldsymbol{\delta}^T \boldsymbol{J}^* (\boldsymbol{\theta})\, \boldsymbol{\delta} = \frac{1}{2} \sum_{\substack{i,j \\ i<j}}^{I} \frac{d_{ij}^2}{p_{ij}} - \sum_{\substack{i,j \\ i<j}} \frac{d_{ij} d_{ji}}{p_{ij}}. \tag{18}$$

Proof. We omit the proof because only requires some algebraic calculations. □

The next theorem is an extension of the Theorems 6 and 8 on page 32 in [9].

Theorem 2. *Let* $\boldsymbol{Z} = (Z_1, \ldots, Z_p)^T$ *be a p-dimensional vector distributed as a normal with mean vector* $\boldsymbol{\mu}_{\mathbf{Z}}$ *and variance–covariance matrix* $\boldsymbol{\Sigma}_{\mathbf{Z}}$. *We assume that* $\boldsymbol{\Sigma}_{\mathbf{Z}}$ *is an idempotent matrix with rank* $l < p$. *Then, for a measurable function* φ, *we have*

(a) $E\left[\boldsymbol{Z}\varphi\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)\right] = \boldsymbol{\mu} E\left[\varphi\left(\chi^2_{p+2}(\lambda)\right)\right]$, *where* $\lambda = E\left[\boldsymbol{Z}^T\boldsymbol{Z}\right]$.

(b) *Let* $\boldsymbol{W}$ *be a positive semi-definite* $I^2 \times I^2$ *matrix. Then,*

$$E\left[\varphi\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)\boldsymbol{Z}^T\boldsymbol{W}\boldsymbol{Z}\right] = \operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{W})\, E\left[\varphi\left(\chi^2_{p+2}(\lambda)\right)\right] + \boldsymbol{\mu}_{\mathbf{Z}}^T \boldsymbol{W} \boldsymbol{\mu}_{\mathbf{Z}} E\left[\varphi\left(\chi^2_{p+4}(\lambda)\right)\right].$$

By $\chi^2_{p+2}(\lambda)$, *we are denoting a non-central chi-square random variable with non-centrality parameter* λ *and by* $\operatorname{tra}(A)$ *the trace of the matrix A.*

Proof. Matrix $\boldsymbol{\Sigma}_{\mathbf{Z}}$ is idempotent with rank $l < I^2$. Therefore there exists an orthogonal matrix $\boldsymbol{C} = (\boldsymbol{C}_1, \boldsymbol{C}_2)$ such that

$$\boldsymbol{C}^T \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{C} = \begin{pmatrix} \mathbf{I}_{l\times l} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{(I^2-l)\times(I^2-l)} \end{pmatrix} \tag{19}$$

and

$$\boldsymbol{C}\boldsymbol{C}^T = \boldsymbol{C}^T \boldsymbol{C} = \mathbf{I}_{I^2\times I^2}. \tag{20}$$

Based on (19), we have

$$\boldsymbol{C}_1^T \boldsymbol{\Sigma}_{\mathbf{Z}} \boldsymbol{C}_1 = \mathbf{I}_{l\times l}$$

and based on (20),

$$\begin{pmatrix} \boldsymbol{C}_1^T\boldsymbol{C}_1 & \boldsymbol{C}_1^T\boldsymbol{C}_2 \\ \boldsymbol{C}_2^T\boldsymbol{C}_1 & \boldsymbol{C}_2^T\boldsymbol{C}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

We define the normal random vector $\boldsymbol{X} = (X_1, \ldots, X_p)^T$ with mean vector $\boldsymbol{C}_1^T\boldsymbol{\mu}_{\mathbf{Z}}$ and variance–covariance matrix $\mathbf{I}_l$. We can write

$$\boldsymbol{Z} = \boldsymbol{C}_1 \boldsymbol{X}.$$

We have $E[\boldsymbol{Z}] = \boldsymbol{C}_1\boldsymbol{C}_1^T\boldsymbol{\mu}_{\mathbf{Z}} = \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{\mu}_{\mathbf{Z}} = \boldsymbol{\mu}_{\mathbf{Z}}$ (the last equality follows because the matrix $\boldsymbol{\Sigma}_{\mathbf{Z}}$ is idempotent and its eigenvalues are 0 or 1), $\operatorname{Var}[\boldsymbol{Z}] = \boldsymbol{C}_1^T\boldsymbol{C}_1 = \boldsymbol{\Sigma}_{\mathbf{Z}}$. The last equality follows because

$$\boldsymbol{C}_1^T\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{C}_1 = \mathbf{I}_m \Longleftrightarrow \boldsymbol{C}_1\boldsymbol{C}_1^T\boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{C}_1 = \boldsymbol{C}_1 \Longleftrightarrow \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{C}_1 = \boldsymbol{C}_1 \Longleftrightarrow \boldsymbol{\Sigma}_{\mathbf{Z}}\boldsymbol{C}_1\boldsymbol{C}_1^T = \boldsymbol{C}_1\boldsymbol{C}_1^T \Longleftrightarrow \boldsymbol{\Sigma}_{\mathbf{Z}} = \boldsymbol{C}_1\boldsymbol{C}_1^T.$$

We can also observe that

$$\boldsymbol{Z}^T\boldsymbol{Z} = \boldsymbol{X}^T\boldsymbol{C}_1^T\boldsymbol{C}_1\boldsymbol{X} = \boldsymbol{X}^T\boldsymbol{X}.$$

Therefore,

$$E\left[\boldsymbol{Z}\varphi\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)\right] = E\left[\boldsymbol{C}_1\boldsymbol{X}\varphi\left(\boldsymbol{X}^T\boldsymbol{X}\right)\right] = \boldsymbol{C}_1 E\left[\boldsymbol{X}\varphi\left(\boldsymbol{X}^T\boldsymbol{X}\right)\right] = \boldsymbol{C}_1\boldsymbol{C}_1^T\boldsymbol{\mu}_{\mathbf{Z}} E\left[\varphi\left(\chi^2_{p+2}(\lambda)\right)\right].$$

The last equality follows by Theorem 6 on page 32 in [9]. Now, we have

$$E\left[\boldsymbol{Z}\varphi\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)\right] = \boldsymbol{\mu}_{\mathbf{Z}} E\left[\varphi\left(\chi^2_{p+2}(\lambda)\right)\right],$$

because $\boldsymbol{C}_1\boldsymbol{C}_1^T = I_{l\times l}$.

Now we are going to establish part (b). We have

$$\begin{aligned} E\left[\varphi\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)\boldsymbol{Z}^T\boldsymbol{W}\boldsymbol{Z}\right] &= E\left[\varphi(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{X}^T\boldsymbol{C}_1^T\boldsymbol{W}\boldsymbol{C}_1\boldsymbol{X}\right] \\ &= \operatorname{tr}\left(\boldsymbol{C}_1^T\boldsymbol{W}\boldsymbol{C}_1\right)E\left[\varphi\left(\chi^2_{p+2}(\lambda)\right)\right] + \boldsymbol{\delta}^T\boldsymbol{\mu}_{\mathbf{Z}}^T\boldsymbol{C}_1\boldsymbol{C}_1^T\boldsymbol{W}\boldsymbol{C}_1\boldsymbol{C}_1^T\boldsymbol{\mu}_{\mathbf{Z}}\boldsymbol{\delta}E\left[\varphi\left(\chi^2_{p+4}(\lambda)\right)\right]. \end{aligned}$$

The last equality follows by Theorem 8 in [9]. Then, we have

$$E\left[\varphi\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)\boldsymbol{Z}^T\boldsymbol{W}\boldsymbol{Z}\right] = \operatorname{tr}\left(\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{W}\right)E\left[h\left(\chi^2_{p+2}(\lambda)\right)\right] + \boldsymbol{\mu}_{\mathbf{Z}}^T\boldsymbol{W}\boldsymbol{\mu}_{\mathbf{Z}}E\left[h\left(\chi^2_{p+4}(\lambda)\right)\right]. \quad \square$$

In Theorems 6 and 8 on page 32 in [9], a similar result was established, but in the particular case that $\boldsymbol{\Sigma}_{\mathbf{Z}} = \mathbf{I}_{p\times p}$.

The *asymptotic bias* of $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, under $H_{1,n}$, is given by

$$\boldsymbol{B}(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})) = \lim_{n\to\infty} E\left[\sqrt{n}\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}_n\right)\right],$$

where $\boldsymbol{p}_n$ was defined in (17). In order to be able to do comparisons, we can consider the *asymptotic quadratic bias* of $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ defined by

$$\operatorname{AQB}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) = \boldsymbol{B}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)^T\boldsymbol{B}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right).$$

In the next theorem we are going to get the expression of $\operatorname{AQB}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$ for the family of *phi-divergence test estimators* $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$.

Theorem 3. *Under $H_{1,n}$, the asymptotic quadratic bias of $\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, defined in (15), is given by*

$$\operatorname{AQB}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) = E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right]^2\lambda, \tag{21}$$

where λ was defined in (18).

Proof. We know that

$$\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) + \left(1 - h\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$$

and

$$\sqrt{n}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}_n\right) = \sqrt{n}\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n\right) - \sqrt{n}\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)h\left(\boldsymbol{Y}_n^T\boldsymbol{Y}_n + o_p(1)\right).$$

Therefore,

$$\begin{aligned} B\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) &= \lim_{n\longrightarrow\infty} E\left[\sqrt{n}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}_n\right)h\left(\boldsymbol{Y}_n^T\boldsymbol{Y}_n + o_p(1)\right)\right] \\ &= \lim_{n\longrightarrow\infty} E\left[\boldsymbol{Y}_n h\left(\boldsymbol{Y}_n^T\boldsymbol{Y}_n + o_p(1)\right)\right] = E\left[\boldsymbol{Y}h\left(\boldsymbol{Y}^T\boldsymbol{Y}\right)\right]. \end{aligned}$$

But $\boldsymbol{Y}$ is a normal random vector with mean vector $\boldsymbol{J}^*(\boldsymbol{\theta})\boldsymbol{\delta}$ and variance–covariance $\boldsymbol{B}(\boldsymbol{\theta})$. It is not difficult to see that $\boldsymbol{B}(\boldsymbol{\theta})$ is an idempotent matrix. Applying part (a) in Theorem 2, we get

$$B\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) = \boldsymbol{J}^*(\boldsymbol{\theta})\boldsymbol{\delta}E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right]$$

and

$$\operatorname{AQB}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) = E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right]^2\boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})^T\boldsymbol{J}^*(\boldsymbol{\theta})\boldsymbol{\delta} = \lambda E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right]^2,$$

because

$$\boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})^T\boldsymbol{J}^*(\boldsymbol{\theta})\boldsymbol{\delta} = \boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\boldsymbol{\delta} = \lambda. \quad \square$$

Based on this theorem, in the next theorem we are going to give some relations between the estimators considered previously.

Theorem 4. *Under $H_{1,n}$, the asymptotic quadratic bias (AQB) of the estimators $\widehat{\boldsymbol{p}}$, $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte}+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{sre}}_{\phi_1}$ and $\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ can be ordered as follows.*

(a) $\operatorname{ADB}(\widehat{\boldsymbol{p}}) \leq \operatorname{ADB}\left(\boldsymbol{p}^{\text{sre}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \operatorname{ADB}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$.

(b) $\mathrm{ADB}\left(\widehat{\boldsymbol{p}}\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$.

(c) $\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$.

(d) $\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$ *iff*

$$E\left[\chi_{m+2}^{-2}(\lambda)\, I_{(0,m-2)}\left(\chi_{m+2}^{2}(\lambda)\right)\right] \geq \frac{1}{(m-2)} G_{m+2}(m-2;\lambda).$$

$G_a(c;\lambda)$ represents the distribution function of a chi-square distribution with a degrees of freedom and non-centrality parameter λ evaluated at the point c.

(e) $\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$ *iff*

$$G_{m+2}\left(\chi_{m,\alpha}^{2};\lambda\right) \geq (m-2)\, E\left[\chi_{m+2}^{-2}(\lambda)\right]$$

for all α and λ.

(f) $\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$ *iff*

$$G_{m+2}\left(\chi_{m+1,\alpha}^{2};\lambda\right) \geq (m-1)E\left[1-\left(1-(m-1)\,\chi_{m+1,\alpha}^{-2}(\lambda)\right) I_{(m-1,\infty)}\left(\chi_{m+1,\alpha}^{2}(\lambda)\right)\right]$$

for all α and λ.

Proof. Based on (21), we have

$$\mathrm{ADB}\left(\widehat{\boldsymbol{p}}\right)=0, \qquad \mathrm{ADB}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=\lambda, \qquad \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=(1-a)^2\lambda,$$

and part (a) follows.

We have

$$\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=\lambda\,(1-a)^2\, G_{m+2}^{2}\left(\chi_{m,\alpha}^{2};\lambda\right) \quad \text{and} \quad \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=\lambda G_{m+2}^{2}\left(\chi_{m,\alpha}^{2};\lambda\right),$$

and part (b) is clear.

From (21), we have

$$\begin{aligned}\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) &= \lambda E\left[1-\left(1-(m-2)\chi_{m+2}^{-2}(\lambda)\right) I_{\left[\chi_{m,\alpha}^{2},\infty\right)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]^2\\ &= \lambda E\left[1-(1-(m-2)\chi_{m+2}^{-2}(\lambda))\left(1-I_{\left(0,\chi_{m,\alpha}^{2}\right)}\left(\chi_{m+2}^{2}(\lambda)\right)\right)\right]^2\\ &= \lambda\left\{G_{m+2}\left(\chi_{m,\alpha}^{2};\lambda\right)+(m-2)\,E\left[\chi_{m+2}^{-2}(\lambda)\right]-(m-2)\,E\left[\chi_{m+2}^{-2}(\lambda)\, I_{\left(0,\chi_{m,\alpha}^{2}\right)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]\right\}^2,\end{aligned}$$

and we know that $\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=\lambda G_{m+2}^{2}\left(\chi_{m,\alpha}^{2};\lambda\right)$. Therefore,

$$\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right) \leq \mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right),$$

because

$$\left\{G_{m+2}\left(\chi_{m,\alpha}^{2};\lambda\right)+(m-2)\,E\left[\chi_{m+2}^{-2}(\lambda)\right]-(m-2)\,E\left[\chi_{m+2}^{-2}(\lambda)\, I_{\left(0,\chi_{m,\alpha}^{2}\right)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]\right\} \geq G_{m+2}\left(\chi_{m,\alpha}^{2};\lambda\right),$$

and part (c) follows.

By (21),

$$\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=(m-2)^2\,\lambda E\left[\chi_{m+2}^{-2}(\lambda)\right]^2$$

and

$$\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)=\lambda\left\{G_{m+2}(m-2;\lambda)+(m-2)\,E\left[\chi_{m+2}^{-2}(\lambda)\right]-(m-2)\,E\left[\chi_{m+2}^{-2}(\lambda)\, I_{(0,m-2)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]\right\}^2.$$

The difference $l=\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)-\mathrm{ADB}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$ can be written as

$$\begin{aligned}l &= \lambda\,(m-2)^2\left\{E\left[\chi_{m+2}^{-2}(\lambda)\, I_{(0,m-2)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]-\frac{1}{(m-2)}G_{m+2}(m-2;\lambda)\right\}\\ &\quad\times\left\{2E\left[\chi_{m+2}^{-2}(\lambda)\right]+\frac{1}{(m-2)}G_{m+2}(m-2;\lambda)-E\left[\chi_{m+2}^{-2}(\lambda)\, I_{(0,m-2)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]\right\},\end{aligned}$$

and it is clear that

$$2E\left[\chi_{m+2}^{-2}(\lambda)\right]+\frac{1}{(m-2)}G_{m+2}(m-2;\lambda)-E\left[\chi_{m+2}^{-2}(\lambda)\,I_{(0,m-2)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]\geq 0$$

and $l\geq 0$ iff $E\left[\chi_{m+2}^{-2}(\lambda)\,I_{(0,m-2)}\left(\chi_{m+2}^{2}(\lambda)\right)\right]\geq\frac{1}{(m-2)}G_{m+2}(m-2;\lambda)$, and part (d) follows.

Parts (e) and (f) follow in a similar way. □

Let $\boldsymbol{W}$ be a positive semi-definite $I^2\times I^2$ matrix. The *weighted square loss function* associated to $\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ is defined by

$$L\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2}),\boldsymbol{p}_n\right)=\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)^T\boldsymbol{W}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right),$$

and the *asymptotic distributional quadratic risk* (ADQR), ADQR $\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)$, of $\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, under $H_{1,n}$, by

$$\lim_{n\to\infty}E\left[\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\theta)}^{-1/2}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\theta)}^{-1/2}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)\right)\right].$$

The following theorem presents the ADQR for the phi-divergence test estimators $\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ defined in (15).

Theorem 5. *Under $H_{1,n}$, the asymptotic distributional quadratic risk (ADQR) of $\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ defined in (15) is given by*

$$\begin{aligned}\text{ADQR}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)&=\operatorname{tr}(\boldsymbol{WJ}(\boldsymbol{\theta}))-\operatorname{tr}(\boldsymbol{WB}(\boldsymbol{\theta}))\left\{2E\left[h\left(\chi_{m+2}^2(\lambda)\right)\right]-E\left[h^2\left(\chi_{m+2}^2(\lambda)\right)\right]\right\}\\&\quad+\boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\boldsymbol{WJ}^*(\boldsymbol{\theta})\boldsymbol{\delta}\left\{2E\left[h\left(\chi_{m+2}^2(\lambda)\right)\right]-2E\left[h\left(\chi_{m+4}^2(\lambda)\right)\right]+E\left[h\left(\chi_{m+4}^2(\lambda)\right)^2\right]\right\}.\end{aligned}$$

Proof. By the definition of the ADQR, and taking into account the expression of $\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, we have

$$\begin{aligned}\text{ADQR}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)&=\lim_{n\to\infty}E\left[\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)\right)\right]\\&=\lim_{n\to\infty}E\left[\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\right)\right]\\&\quad-2\lim_{n\to\infty}E\left[\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}h\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\left(\widehat{\boldsymbol{p}}-\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\right)\right]\\&\quad+\lim_{n\to\infty}E\left[h\left(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)^2\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\widehat{\boldsymbol{p}}-\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\widehat{\boldsymbol{p}}-\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)\right].\end{aligned}$$

It is well known that if $\boldsymbol{A}$ is a symmetric nonnegative definite matrix and $\boldsymbol{Y}$ an M-dimensional random vector with mean vector $\boldsymbol{\mu}$ and variance–covariance matrix $\boldsymbol{\Sigma}$,

$$E\left[(\boldsymbol{Y}-\boldsymbol{a})^T\boldsymbol{A}(\boldsymbol{Y}-\boldsymbol{a})\right]=\operatorname{tr}(\boldsymbol{A\Sigma})+(\boldsymbol{\mu}-\boldsymbol{a})^T\boldsymbol{A}(\boldsymbol{\mu}-\boldsymbol{a}).$$

Applying this result, we get

$$\lim_{n\to\infty}E\left[\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\right)\right]=\operatorname{tr}(\boldsymbol{WJ}(\boldsymbol{\theta})),$$

because

$$\sqrt{n}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\underset{n\to\infty}{\overset{L}{\to}}N\left(\boldsymbol{0},\boldsymbol{\Sigma}_{\boldsymbol{p}(\boldsymbol{\theta})}\right)$$

and

$$\sqrt{n}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}(\widehat{\boldsymbol{p}}-\boldsymbol{p}_n)\underset{n\to\infty}{\overset{L}{\to}}N\left(\boldsymbol{0},\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\boldsymbol{\Sigma}_{\boldsymbol{p}(\boldsymbol{\theta})}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\right),$$

with $\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\boldsymbol{\Sigma}_{\boldsymbol{p}(\boldsymbol{\theta})}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}=\boldsymbol{J}(\boldsymbol{\theta})$.

On the other hand,

$$\lim_{n\to\infty}E\left[\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}h(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}))^2\left(\widehat{\boldsymbol{p}}-\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)^T\boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\widehat{\boldsymbol{p}}-\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)\right]=E\left[h(\boldsymbol{Y}^T\boldsymbol{Y})^2\boldsymbol{Y}\boldsymbol{W}\boldsymbol{Y}\right],$$

and part (b) in Theorem 2 gives

$$E\left[h(\boldsymbol{Y}^T\boldsymbol{Y})^2\boldsymbol{Y}^T\boldsymbol{W}\boldsymbol{Y}\right] = \operatorname{tr}\left(\boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{W}\right) E\left[h\left(\chi^2_{m+2}(\lambda)\right)^2\right] + \boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta} E\left[h\left(\chi^2_{m+4}(\lambda)\right)^2\right].$$

Now, we get

$$l = \lim_{n\to\infty} E\left[\sqrt{n}\left(\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})} h(T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}))\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)\right)^T \boldsymbol{W}\sqrt{n}\left(\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n)\right)\right].$$

We denote $\boldsymbol{X}_n^* = \boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\sqrt{n}(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n)$, and we are going to get the joint distribution of the random vector $\boldsymbol{X}_n^*$ and $\boldsymbol{Y}_n = \sqrt{n}\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$.

It is well known (see Chapter 6 in [13]) that the minimum phi-divergence estimator, $\widehat{\boldsymbol{\beta}}^{\phi_2}$, verifies

$$\widehat{\boldsymbol{\beta}}^{\phi_2} = \boldsymbol{\beta} + \boldsymbol{I}_F(\boldsymbol{\beta})^{-1}\boldsymbol{A}(\boldsymbol{\beta})^T\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))\right) + o\left(\|\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))\|\right), \tag{22}$$

where

$$\boldsymbol{A}(\boldsymbol{\beta}) = \boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))}\frac{\partial \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))}{\partial\boldsymbol{\beta}} = \boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))}\frac{\partial \boldsymbol{p}(\boldsymbol{g}(\boldsymbol{\beta}))}{\partial\boldsymbol{g}(\boldsymbol{\beta})}\left(\frac{\partial\boldsymbol{g}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right)^T.$$

On the other hand,

$$\boldsymbol{g}\left(\widehat{\boldsymbol{\beta}}^{\phi_2}\right) = \boldsymbol{g}(\boldsymbol{\beta}) + \left(\frac{\partial(\boldsymbol{g}(\boldsymbol{\beta}))}{\partial\boldsymbol{\beta}}\right)^T\left(\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right) + o\left(\left\|\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right\|\right).$$

But $\boldsymbol{g}\left(\widehat{\boldsymbol{\beta}}^{\phi_2}\right) = \widehat{\boldsymbol{\theta}}^{\phi_2}$, $\boldsymbol{g}(\boldsymbol{\beta}) = \boldsymbol{\theta}$, and the expression of $\left(\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right)$ can be obtained from (22). Therefore,

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}^{\phi_2} &= \boldsymbol{\theta} + \left(\frac{\partial(\boldsymbol{g}(\boldsymbol{\beta}))}{\partial\boldsymbol{\beta}}\right)^T\boldsymbol{I}_F^S(\boldsymbol{\beta})^{-1}\frac{\partial(\boldsymbol{g}(\boldsymbol{\beta}))}{\partial\boldsymbol{\beta}}\boldsymbol{A}(\boldsymbol{\theta})^T\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\right) + o\left(\|\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\|\right) + o\left(\left\|\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right\|\right)\\
&= \boldsymbol{\theta} + \left(\frac{\partial\boldsymbol{g}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right)^T\boldsymbol{I}_F(\boldsymbol{\beta})^{-1}\frac{\partial\boldsymbol{g}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\left(\frac{\partial\boldsymbol{p}\boldsymbol{g}(\boldsymbol{\beta})}{\partial\boldsymbol{g}(\boldsymbol{\beta})}\right)^T\boldsymbol{D}^{-1}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\right) + o\left(\|\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\|\right) + o\left(\left\|\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right\|\right)\\
&= \boldsymbol{\theta} + \boldsymbol{I}_F(\boldsymbol{\theta})^{-1}\boldsymbol{A}(\boldsymbol{\theta})^T\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\right) + o\left(\|\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\|\right) + o\left(\left\|\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right\|\right).
\end{aligned}$$

Then,

$$\begin{aligned}
\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}(\boldsymbol{\theta}) &= \boldsymbol{T}\left(\widehat{\boldsymbol{\theta}}^{\phi_2} - \boldsymbol{\theta}\right) + o\left(\left\|\widehat{\boldsymbol{\theta}}^{\phi_2} - \boldsymbol{\theta}\right\|\right)\\
&= \boldsymbol{T}\boldsymbol{M}_{\boldsymbol{\beta}}^T\boldsymbol{I}_F(\boldsymbol{\beta})^{-1}\boldsymbol{M}_{\boldsymbol{\beta}}\boldsymbol{A}(\boldsymbol{\theta})^T\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\right) + o\left(\|\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\|\right) + o\left(\left\|\widehat{\boldsymbol{\beta}}^{\phi_2} - \boldsymbol{\beta}\right\|\right).
\end{aligned}$$

Defining

$$\boldsymbol{L}(\boldsymbol{\theta}) = \boldsymbol{T}\boldsymbol{M}_{\boldsymbol{\beta}}^T\boldsymbol{I}_F(\boldsymbol{\beta})^{-1}\boldsymbol{M}_{\boldsymbol{\beta}}\boldsymbol{A}(\boldsymbol{\theta})^T\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})},$$

we have

$$\sqrt{n}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}(\boldsymbol{\theta})\right) = \sqrt{n}\boldsymbol{L}(\boldsymbol{\theta})\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n\right) + \boldsymbol{L}(\boldsymbol{\theta})\boldsymbol{d} + o_p(1).$$

Now, we get

$$\sqrt{n}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})\right) = \sqrt{n}\boldsymbol{I}_{I^2}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n\right) + \boldsymbol{d} + o_p(1)$$

and

$$\sqrt{n}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}(\boldsymbol{\theta})\right) = \sqrt{n}\boldsymbol{L}(\boldsymbol{\theta})\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n\right) + \boldsymbol{L}(\boldsymbol{\theta})\boldsymbol{d} + o_p(1).$$

Then, the random vector

$$\sqrt{n}\boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})\right)$$

is asymptotically normal, with mean vector

$$\boldsymbol{\mu} = \boldsymbol{D}^{-1/2}_{\boldsymbol{p}(\boldsymbol{\theta})}\left(\boldsymbol{I}_{I^2\times I^2} - \boldsymbol{L}(\boldsymbol{\theta})\right)\boldsymbol{d},$$

and variance–covariance matrix

$$\boldsymbol{C}(\boldsymbol{\theta}) \equiv \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}} - \boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{L}(\boldsymbol{\theta})^T - \boldsymbol{L}(\boldsymbol{\theta})\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}} + \boldsymbol{L}(\boldsymbol{\theta})\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{L}(\boldsymbol{\theta})^T\right)\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}.$$

It is not difficult to establish that $\boldsymbol{\mu} = \boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}$, $\boldsymbol{C}(\boldsymbol{\theta}) = \boldsymbol{B}(\boldsymbol{\theta})$ and $\boldsymbol{B}(\boldsymbol{\theta})$ is an idempotent matrix with rank$(\boldsymbol{B}(\boldsymbol{\theta})) = m$. If we denote $\boldsymbol{S}_n = \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\sqrt{n}\,(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n)$, we can write

$$\begin{pmatrix}\boldsymbol{S}_n\\ \boldsymbol{Y}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\\ \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(I_{I^2\times I^2} - \boldsymbol{L}(\boldsymbol{\theta})\right)\end{pmatrix}\sqrt{n}\,(\widehat{\boldsymbol{p}} - \boldsymbol{p}_n) + \begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{D}_{\boldsymbol{p}(\boldsymbol{\theta})}^{-1/2}\left(I_{I^2\times I^2} - \boldsymbol{L}(\boldsymbol{\theta})\right)\boldsymbol{d}\end{pmatrix} + \begin{pmatrix}o_p(1)\\ o_p(1)\end{pmatrix}$$

and

$$\begin{pmatrix}\boldsymbol{S}_n\\ \boldsymbol{Y}_n\end{pmatrix} \underset{n\to\infty}{\overset{L}{\to}} \mathcal{N}\left(\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}\end{pmatrix}, \begin{pmatrix}\boldsymbol{J}(\boldsymbol{\theta}) & \boldsymbol{B}(\boldsymbol{\theta})\\ \boldsymbol{B}(\boldsymbol{\theta}) & \boldsymbol{B}(\boldsymbol{\theta})\end{pmatrix}\right).$$

Let $\boldsymbol{S}$ be the random vector verifying

$$\boldsymbol{S}_n \underset{n\to\infty}{\overset{L}{\to}} \boldsymbol{S}.$$

We have

$$E\left[\boldsymbol{S}/\boldsymbol{Y} = \boldsymbol{y}\right] = \boldsymbol{y} - \boldsymbol{J}^*(\boldsymbol{\theta}_0)\,\boldsymbol{\delta}. \tag{23}$$

Based on this result, we get

$$l = E\left[\boldsymbol{Y}^T\boldsymbol{W}\boldsymbol{S}h\left(\boldsymbol{Y}^T\boldsymbol{Y}\right)\right] = E\left[h\left(\boldsymbol{Y}^T\boldsymbol{Y}\right)\boldsymbol{Y}^T\boldsymbol{W}E\left[\boldsymbol{S}/\boldsymbol{Y}\right]\right].$$

Now, by (23), we can write

$$l = E\left[h\left(\boldsymbol{Y}^T\boldsymbol{Y}\right)\boldsymbol{Y}^T\boldsymbol{W}\boldsymbol{Y}\right] - E\left[h\left(\boldsymbol{Y}^T\boldsymbol{Y}\right)\boldsymbol{Y}^T\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}\right].$$

Applying Theorem 2, we get

$$l = \operatorname{tr}\left(\boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{W}\right)E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right] + \boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}E\left[h\left(\chi^2_{m+4}(\lambda)\right)\right] - \boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right].$$

Now the result follows, because

$$\begin{aligned}\operatorname{ADQR}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) = {} & \operatorname{tr}(\boldsymbol{W}\boldsymbol{J}(\boldsymbol{\theta})) + \operatorname{tr}\left(\boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{W}\right)E\left[h\left(\chi^2_{m+2}(\lambda)\right)^2\right] + \boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}E\left[h\left(\chi^2_{m+4}(\lambda)\right)^2\right]\\ & - 2\operatorname{tr}\left(\boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{W}\right)E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right] - 2\boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}E\left[h\left(\chi^2_{m+4}(\lambda)\right)\right]\\ & + 2\boldsymbol{\delta}^T\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{W}\boldsymbol{J}^*(\boldsymbol{\theta})\,\boldsymbol{\delta}E\left[h\left(\chi^2_{m+2}(\lambda)\right)\right]. \quad \square\end{aligned}$$

In the following theorem, we present, under the null hypothesis of symmetry, some relations between the estimators $\widehat{\boldsymbol{p}}$, $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte+}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ and $\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$.

Theorem 6. *Under the null hypothesis of symmetry and weighted square loss function*

$$L\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}), \boldsymbol{p}_n\right) = \left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}_n\right)^T\boldsymbol{W}\left(\boldsymbol{p}^h_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}) - \boldsymbol{p}_n\right)$$

with positive semi-definite matrix $\boldsymbol{W}$, *the asymptotic distributional quadratic risk of the estimators* $\widehat{\boldsymbol{p}}$, $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte+}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ *and* $\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ *can be ordered as follows:*

(a) $\operatorname{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}\left(\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}\left(\boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}(\widehat{\boldsymbol{p}};\boldsymbol{W})$.

(b) *If* $m > 2$,

$$\operatorname{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}\left(\boldsymbol{p}^{S+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}\left(\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}(\widehat{\boldsymbol{p}};\boldsymbol{W})$$

and

$$\operatorname{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}\left(\boldsymbol{p}^{\text{pte+}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}\left(\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) \le \operatorname{ADQR}(\widehat{\boldsymbol{p}};\boldsymbol{W}).$$

(c) $\operatorname{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right) < \operatorname{ADQR}\left(\widehat{\boldsymbol{\beta}}^{\text{sre}}_{\phi_2};\boldsymbol{W}\right) < \operatorname{ADQR}(\widehat{\boldsymbol{p}};\boldsymbol{W})$.

(d) *If* $G_{m+2}(\chi^2_{m,\alpha}; 0) \geq (m-2)/2$,

$$R\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}(\widehat{\boldsymbol{p}}; \boldsymbol{W}).$$

Proof. Under the hypothesis of symmetry $\lambda = 0$, and, by Theorem 5, we have

$$\text{ADQR}\left(\boldsymbol{p}^{h}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = \text{tr}\left(\boldsymbol{WK}(\boldsymbol{\theta})\right) + \text{tr}\left(\boldsymbol{WB}(\boldsymbol{\theta})\right) E\left[\left(1 - h\left(\chi^2_{m+2}(0)\right)\right)^2\right].$$

Therefore the ADQR is an increasing function of

$$E\left[\left(1 - h\left(\chi^2_{m+2}(0)\right)\right)^2\right]. \tag{24}$$

In the following table, we present the expressions of $E[(1-h(\chi^2_{m+2}(0)))^2]$ for the different estimators $\widehat{\boldsymbol{p}}$, $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{spt}}_{\phi_1}, (\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, $\boldsymbol{p}^{\text{pte}+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ and $\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$.

Estimator	$E\left[\left(1 - h\left(\chi^2_{m+2}(0)\right)\right)^2\right]$
$\widehat{\boldsymbol{p}}$	1
$\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	0
$\boldsymbol{p}^{\text{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	a^2
$\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	$1 - G_{m+2}(\chi^2_{m,\alpha}; 0)$
$\boldsymbol{p}^{\text{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	$1 - a(2-a)\, G_{m+2}(\chi^2_{m,\alpha}; 0)$
$\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	$\int_0^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0) = 1 - \frac{m-2}{m}$
$\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	$\int_{m-2}^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0)$
$\boldsymbol{p}^{\text{pte}+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$	$\int_{\chi^2_{r,\alpha}}^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0)$

Expressions of $E\left[\left(1 - h\left(\chi^2_{m+2}(0)\right)\right)^2\right]$.

Based on this table, it is immediate that

$$\text{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\boldsymbol{p}^{\text{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\boldsymbol{p}^{\text{spt}}_{\phi_1}; \boldsymbol{W}\right) \leq \text{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right).$$

We consider $m > 2$. If X is a Gamma random variable with parameters p and a, we have

$$E\left[X^{-1}\right] = \frac{a}{p-1} \quad \text{and} \quad E\left[X^{-2}\right] = \frac{a^2}{(p-1)(p-2)};$$

therefore,

$$\int_0^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0) = 1 + (m-2)^2 \frac{1/4}{\left(\frac{m+2}{2} - 1\right)\left(\frac{m+2}{2} - 2\right)} - 2(m-2)\frac{1/2}{\frac{m+2}{2} - 1} = 1 - \frac{m-2}{m}.$$

Expression (24) for $\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ is

$$\int_0^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0) = 1 - \frac{m-2}{m}$$

and for $\widehat{\boldsymbol{p}}$ it is 1. Therefore, we have

$$\text{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \text{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right).$$

The inequality between $\text{ADQR}\left(\boldsymbol{p}^{s+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ and $\text{ADQR}\left(\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ follows because

$$\int_{m-2}^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0) \leq \int_0^\infty (1-(m-2)x^{-1})^2 \text{d}G_{m+2}(x; 0).$$

We have

$$\int_{\chi^2_{m,\alpha}}^{\infty}(1-(m-2)x^{-1})^2 \mathrm{d}G_{m+2}(x;0) = 1-G_{m+2}(\chi^2_{m,\alpha};0)+\frac{m-2}{m}\left\{-\left(1-G_m(\chi^2_{m,\alpha};0)\right)\right.$$
$$\left.-\left(G_{m-2}(\chi^2_{m,\alpha};0)-G_m(\chi^2_{m,\alpha};0)\right)\right\}$$

and

$$\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)-\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)$$
$$=\frac{m-2}{m}\left\{-\left(1-G_m(\chi^2_{m,\alpha};0)\right)-\left(G_{m-2}(\chi^2_{m,\alpha};0)-G_m(\chi^2_{m,\alpha};0)\right)\right\}.$$

Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)$ because $G_{m-2}(\chi^2_{m,\alpha};0)>G_m(\chi^2_{m,\alpha};0)$.

Now, it is clear that

$$\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}+}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right).$$

It is also immediate to see that $\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)<\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{spt}}_{\phi_1};\boldsymbol{W}\right)<\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right)$.

Finally to see (d), we have

$$1-G_{m+2}(\chi^2_{m,\alpha};0)=\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\boldsymbol{p}^{s}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)$$
$$=\int_0^{\infty}(1-(m-2)x^{-1})^2\mathrm{d}G_{m+2}(x;0)=1-\frac{m-2}{m};$$

the inequality is true iff $G_{m+2}(\chi^2_{m,\alpha};0)\ge\frac{m-2}{m}$. $\square$

In the following theorem the following result (Courant Theorem) will be important. Let $\lambda_1,\ldots,\lambda_n$ be the characteristic roots of an $n\times n$ matrix $\boldsymbol{A}$. Then,

$$\mathrm{Ch}_{\min}(\boldsymbol{A})\le\frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}\le\mathrm{Ch}_{\max}(\boldsymbol{A}),$$

with $\mathrm{Ch}_{\min}(\boldsymbol{A})=\min\{\lambda_1,\ldots,\lambda_n\}$ and $\mathrm{Ch}_{\max}(\boldsymbol{A})=\max\{\lambda_1,\ldots,\lambda_n\}$.

Theorem 7. *Under the contiguous alternative hypothesis, given in (17), and the weighted square loss function*

$$L\left(\boldsymbol{p}^{h}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2}),\boldsymbol{p}_n\right)=\left(\boldsymbol{p}^{h}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)^T\boldsymbol{W}\left(\boldsymbol{p}^{h}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})-\boldsymbol{p}_n\right)$$

with positive semi-definite matrix $\boldsymbol{W}$, we have the following results in relation to the ADQR.

(a) $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right)$ *iff*

$$\lambda\le\frac{\mathrm{tr}\left(\boldsymbol{W}\boldsymbol{B}(\boldsymbol{\theta})\right)\left(1-a^2\right)}{(1-a)^2\,\mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

(b) $\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right)$ *iff*

$$\lambda\le\frac{\mathrm{tr}\left(\boldsymbol{W}\boldsymbol{B}(\boldsymbol{\theta})\right)}{\mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

(c) $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right)$ *iff*

$$\lambda\le\frac{\mathrm{tr}\left(\boldsymbol{W}\boldsymbol{B}(\boldsymbol{\theta})\right)G_{m+2}\left(\chi^2_{m;\alpha};\lambda\right)}{\left(2G_{m+2}\left(\chi^2_{m;\alpha};\lambda\right)-G_{m+4}\left(\chi^2_{m;\alpha};\lambda\right)\right)\mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

If $\lambda\to\infty$ or $\alpha\to1$ we have $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{pte}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\to\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right)$.

(d) $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{spt}}_{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2});\boldsymbol{W}\right)\le\mathrm{ADQR}\left(\widehat{\boldsymbol{p}};\boldsymbol{W}\right)$ *iff*

$$\lambda\le\frac{\left(1-a^2\right)\ \mathrm{tr}\left(\boldsymbol{W}\boldsymbol{B}(\boldsymbol{\theta})\right)G_{m+2}\left(\chi^2_{m;\alpha};\lambda\right)}{\left\{\left(a^2-1\right)G_{m+4}\left(\chi^2_{m;\alpha};\lambda\right)+2\left(1-a\right)G_{m+2}\left(\chi^2_{m;\alpha};\lambda\right)\right\}\mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

(e) $\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ *iff*

$$\lambda \le \frac{\mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\left(1 - G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right)}{\left\{1 - 2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) + G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}\mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

(f) $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ *iff*

$$\lambda \le \frac{(1-a)\,G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)}{\left\{2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) - (1-a)G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}\mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

(g) *Let* $m > 2$, *and assume that*

$$\mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\,\mathrm{Ch}_{\max}(\boldsymbol{W})^{-1} \ge \frac{m+2}{2}; \tag{25}$$

then, for all λ,

$$\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right).$$

If, in addition,

$$\boldsymbol{\delta}^T \boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{WJ}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta} > \mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta})),$$

we have

$$\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \ge \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right).$$

(h) *Let* $0 < a < 1$. *We have*

$$\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right).$$

Proof. By Theorem 5, we have

$$\mathrm{ADQR}\,(\widehat{\boldsymbol{p}}; \boldsymbol{W}) = \mathrm{tr}\,(\boldsymbol{WJ}\,(\boldsymbol{\theta}))$$

and

$$\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = \mathrm{tr}\,(\boldsymbol{WJ}\,(\boldsymbol{\theta})) - \mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\left(1-a^2\right) + (1-a)^2\,\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{WJ}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}.$$

Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\,(\widehat{\boldsymbol{p}}; \boldsymbol{W})$ iff

$$(1-a)^2\,\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{WJ}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta} \le \mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\left(1-a^2\right).$$

This is equivalent to

$$\frac{\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{WJ}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}}{\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}} \le \frac{\mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\left(1-a^2\right)}{\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}\,(1-a)^2}.$$

But $\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta} = \boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta} = \lambda$; then

$$\frac{\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{WJ}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}}{\lambda} \le \frac{\mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\left(1-a^2\right)}{\lambda\,(1-a)^2}.$$

Applying the Courant Theorem,

$$\mathrm{Ch}_{\min}(\boldsymbol{W}) \le \frac{\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{WJ}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}}{\boldsymbol{\delta}^T\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{J}^*\,(\boldsymbol{\theta})\,\boldsymbol{\delta}} \le \mathrm{Ch}_{\max}(\boldsymbol{W}).$$

Therefore $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \le \mathrm{ADQR}\,(\widehat{\boldsymbol{p}}; \boldsymbol{W})$ iff

$$\lambda \le \frac{\mathrm{tr}\,(\boldsymbol{WB}\,(\boldsymbol{\theta}))\left(1-a^2\right)}{(1-a)^2\,\mathrm{Ch}_{\max}(\boldsymbol{W})}$$

and $\mathrm{ADQR}\left(\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \geq \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right)$ iff $\lambda \geq \frac{\mathrm{tr}(\boldsymbol{WB}(\boldsymbol{\theta}))\left(1-a^2\right)}{(1-a)^2 \mathrm{Ch}_{\max}(\boldsymbol{W})}$. Result (a) follows. Result (b) follows because $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ can be obtained from $\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ with $a = 0$. We know that

$$\begin{aligned}\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = {} & \mathrm{tr}\left(\boldsymbol{WJ}\left(\boldsymbol{\theta}\right)\right) - \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) \\ & + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) - G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}.\end{aligned}$$

If $\lambda \to \infty$, then $G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) \to 0$ and $G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) \to 0$. Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \to \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right)$. Similarly, if $\alpha \to 1$, then $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \to \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right)$.

Now $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right)$ iff

$$\frac{\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)}{\lambda} \geq \frac{\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}}{\lambda}\left\{2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) - G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\},$$

and this is equivalent to

$$\lambda \leq \frac{\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)}{\mathrm{Ch}_{\max}(\boldsymbol{W})\left\{2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) - G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}}.$$

Now we are going to establish (d). We have

$$\begin{aligned}\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = {} & \mathrm{tr}\left(\boldsymbol{WJ}\left(\boldsymbol{\theta}\right)\right) + \left(a^2 - 1\right) \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) \\ & + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{\left(a^2 - 1\right) G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2\left(1 - a\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}.\end{aligned}$$

Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right)$ iff

$$\left(a^2 - 1\right) \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{\left(a^2 - 1\right) G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2\left(1 - a\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\} \leq 0,$$

or, equivalently,

$$\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{\left(a^2 - 1\right) G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2\left(1 - a\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\} \leq \left(1 - a^2\right) \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right).$$

Finally, we have $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \mathrm{ADQR}\left(\widehat{\boldsymbol{p}}; \boldsymbol{W}\right)$ iff

$$\lambda \leq \frac{\left(1 - a^2\right) \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)}{\left\{\left(a^2 - 1\right) G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2\left(1 - a\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\} \mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

No we are going to see (e). We know that

$$\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = \mathrm{tr}\left(\boldsymbol{WJ}\left(\boldsymbol{\theta}\right)\right) - \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}$$

and

$$\begin{aligned}\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = {} & \mathrm{tr}\left(\boldsymbol{WJ}\left(\boldsymbol{\theta}\right)\right) - \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) \\ & + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) - G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}.\end{aligned}$$

Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ iff

$$\frac{\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}}{\lambda} \leq \frac{\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right)\left\{1 - G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}}{\left\{1 - 2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) + G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}};$$

but

$$\frac{\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}}{\lambda} \leq \mathrm{Ch}_{\max}(\boldsymbol{W}),$$

and the result follows.

Part (f). We have

$$\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = \mathrm{tr}\left(\boldsymbol{WJ}\left(\boldsymbol{\theta}\right)\right) + \left(a^2-1\right)\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{\left(a^2-1\right) G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2\left(1-a\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}$$

and

$$\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = \mathrm{tr}\left(\boldsymbol{WJ}\left(\boldsymbol{\theta}\right)\right) - \mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) + \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{2G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) - G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}.$$

Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\mathrm{pte}}); \boldsymbol{W}\right) \leq \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{\mathrm{spt}}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ iff

$$\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left\{-a^2 G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2aG_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\} \leq a^2\,\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right).$$

Taking into account that $0 < a < 1$ and $G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) \leq G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)$, we have $-a^2 G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2aG_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right) \geq 0$. Then

$$\frac{\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}}{\lambda} \leq \frac{a^2\,\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)}{\lambda\left\{-a^2 G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2aG_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\}},$$

and based on this inequality, and the Courant Theorem, we have

$$\lambda \leq \frac{a^2\,\mathrm{tr}\left(\boldsymbol{WB}\left(\boldsymbol{\theta}\right)\right) G_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)}{\left\{-a^2 G_{m+4}\left(\chi^2_{m;\alpha}; \lambda\right) + 2aG_{m+2}\left(\chi^2_{m;\alpha}; \lambda\right)\right\} \mathrm{Ch}_{\max}(\boldsymbol{W})}.$$

From Theorem 5, and denoting $t = \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) - \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$, we have

$$\begin{aligned} t &= -\mathrm{tr}\left(\boldsymbol{\Sigma}_{\mathbf{Y}}\boldsymbol{M}\right) E\left[\left(1-(m-2)\,\chi^{-2}_{m+2}\left(\lambda\right)\right)^2 I_{(0,m-2)}(\chi^2_{m+2}(\lambda))\right] \\ &\quad - \boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta} E\left[\left(1-(m-2)\,\chi^{-2}_{m+4}\left(\lambda\right)\right)^2 I_{(0,m-2)}(\chi^2_{m+4}(\lambda))\right] \\ &\quad - 2\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta} E\left[\left((m-2)\,\chi^{-2}_{m+2}\left(\lambda\right)-1\right) I_{(0,m-2)}(\chi^2_{m+2}(\lambda))\right]. \end{aligned}$$

Therefore, $\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s+}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) \leq \mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right)$ because

$$E\left[\left(1-(m-2)\,\chi^{-2}_{m+2}\left(\lambda\right)\right) I_{(0,m-2)}\left(\chi^2_{m+2}\left(\lambda\right)\right)\right] = \int_0^{m-2}\left(1-(m-2)\,x^{-1}\right) \mathrm{d}G_{m+2}(x;\lambda) < 0.$$

By Courant's Theorem it is a simple exercise to establish that $\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta} \leq \lambda \mathrm{Ch}_{\max}\left(\boldsymbol{W}\right)$.

By Theorem 5, we get

$$\mathrm{ADQR}\left(\boldsymbol{p}_{\phi_1}^{s}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) - R\left(\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2}); \boldsymbol{W}\right) = -\left(m-2\right)\mathrm{tr}(\boldsymbol{M}\boldsymbol{\Sigma}_{\mathbf{Y}})\left\{\left(m-2\right) E\left[\chi^{-4}_{m+2}\left(\lambda\right)\right] + \left[1-\frac{\boldsymbol{\delta}^T \boldsymbol{J}^*\left(\boldsymbol{\theta}\right) \boldsymbol{WJ}^*\left(\boldsymbol{\theta}\right) \boldsymbol{\delta}\left(m+2\right)}{2\lambda\mathrm{tr}(\boldsymbol{M}\boldsymbol{\Sigma}_{\mathbf{Y}})}\right] 2\lambda E\left[\chi^{-4}_{m+4}\left(\lambda\right)\right]\right\} \leq 0.$$

The last inequality follows by (25).

Part (h) is immediate. □

4. A numerical application

In order to clarify the different preliminary phi-divergence test estimators, $\boldsymbol{p}_{\phi_1}^{h}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, introduced and studied in this paper for the symmetry model, in this section we are going to consider a numerical example.

We shall consider the data in Table 1.

These data were collected by Glass [16] in a study of social mobility in Great Britain. They are a cross-classification of a sample of British males according to each subject's status category and his father's status category. A question of interest might be whether or not changes in class between fathers and sons occur in both directions with the same probability, i.e., to determine whether observations in cells situated symmetrically about the main diagonal have the same probability

Table 1
Cross-classification of a sample of British males to each subject's status category and his father's status category.

		Subject's status		
		Upper	Midle	Lower
Father's status	Upper	588	395	159
	Midle	349	714	447
	Lower	114	320	411

of occurrence, i.e., if $p_{ij} = p_{ji}$ $(i \neq j)$. Under such a hypothesis, the frequencies in the ijth and jith cells are expected to be equal, and the maximum likelihood estimator is

$$\boldsymbol{p}\left(\widehat{\boldsymbol{\theta}}_{ij}\right) = \begin{cases} \frac{1}{2}\left(n_{ij}+n_{ji}\right) & i \neq j \\ n_{ij} & i = j, \end{cases} \qquad i = 1, 2, 3.$$

When the hypothesis of symmetry holds, $\boldsymbol{p}\left(\widehat{\boldsymbol{\theta}}\right) = \left(\boldsymbol{p}\left(\widehat{\boldsymbol{\theta}}_{11}\right), \dots, \boldsymbol{p}\left(\widehat{\boldsymbol{\theta}}_{33}\right)\right)$ has a smaller risk with a quadratic loss than the unrestricted estimators $\widehat{\boldsymbol{p}} = \left(\frac{n_{ii}}{n}, \dots, \frac{n_{33}}{n}\right)$, where n_{ij} is the number of elements in the cell $(1, 1)$ and n is the number of elements in the sample $(n = n_{11} + \cdots + n_{33})$. If the hypothesis of symmetry does not hold, the risk of $\boldsymbol{p}\left(\widehat{\boldsymbol{\theta}}\right)$ may be very big. For this reason, the prior knowledge about the hypothesis of symmetry $\left(p_{ij} = p_{ji}\right)$ is rather uncertain, and it may be desirable to use a preliminary test estimator. We are going to obtain the different estimators considered in this paper, for the data in Table 1.

If we observe the expression of the preliminary test estimators in (9), we can see that we need previously to get $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ (the minimum ϕ_2-divergence estimator for the probability vector $\boldsymbol{p} = (p_{11}, \dots, p_{33})$) as well as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ (the family of test statistics for testing $p_{ij} = p_{ji}$, $i, j = 1, 2, 3$).

In order to obtain $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{\phi_2})$ as well as $T_n^{\phi_1}(\widehat{\boldsymbol{\theta}}^{\phi_2})$, we shall consider

$$\phi_2(x) = \phi_1(x) = \phi_\lambda(x) = \frac{1}{\lambda(\lambda+1)}\left(x^{\lambda+1} - x - \lambda(x-1)\right) \tag{26}$$

for $\lambda \neq 0$ and $\lambda \neq -1$ and

$$\phi_0(x) = \lim_{\lambda \longrightarrow 0} \phi_\lambda(x) = x \log x - x + 1$$
$$\phi_{-1}(x) = \lim_{\lambda \longrightarrow -1} \phi_\lambda(x) = -\log x + x + 1.$$

In the following, we denote by $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ the family of minimum divergence estimators associated to $\phi_2(x)$, given in (26), and by $T_n^{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ the family of test statistics associated to $\phi_2(x)$ and $\phi_1(x)$, given in (26), and defined in (8).

The expression of $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$, see for instance page 371 in [13], is given by

$$\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) = \frac{\left(\frac{\widehat{p}_{ij}^{\lambda_2+1} + \widehat{p}_{ji}^{\lambda_2+1}}{2}\right)^{\frac{1}{\lambda_2+1}}}{\sum_{i=1}^{3}\sum_{j=1}^{3}\left(\frac{\widehat{p}_{ij}^{\lambda_2+1} + \widehat{p}_{ji}^{\lambda_2+1}}{2}\right)^{\frac{1}{\lambda_2+1}}}, \quad i, j = 1, \dots, 3 : \lambda \neq 0,\ \lambda \neq -1, \tag{27}$$

and the expression of $T_n^{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ by

$$T_n^{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) = \frac{1}{\lambda_1(1+\lambda_1)}\left(\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{p_{ij}^{\lambda_1+1}}{\left(\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})\right)^{\lambda_1}} - 1\right), \quad \lambda \neq 0,\ \lambda \neq -1. \tag{28}$$

Important and well-known estimators obtained from (27) are the following.

For $\lambda_2 = 0$, we obtain the maximum likelihood estimator for symmetry introduced by Bowker [1], whose expression is

$$\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}^{(0)}) = \frac{\widehat{p}_{ij} + \widehat{p}_{ji}}{2}, \quad i, j = 1, \dots, 3.$$

For $\lambda = -1$, we obtain, as a limit case,

$$\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}^{(-1)}) = \frac{\left(\widehat{p}_{ij}\widehat{p}_{ji}\right)^{\frac{1}{2}}}{\sum_{i=1}^{3}\sum_{j=1}^{3}\left(\widehat{p}_{ij}\widehat{p}_{ji}\right)^{\frac{1}{2}}}, \quad i, j = 1, \dots, 3$$

Table 2
Minimum ϕ_2-divergence estimators for different values of λ_2 (− 1, −1/2, 0, 1/2, 1, 2/3).

Cell	$\lambda_2 = -1$	$\lambda_2 = -1/2$	$\lambda_2 = 0$	$\lambda_2 = 1/2$	$\lambda_2 = 1$	$\lambda_2 = 2/3$
11	0.16924	0.16852	0.16814	0.16765	0.16740	0.16765
12	0.10687	0.10651	0.10638	0.10624	0.10611	0.10620
13	0.03875	0.03885	0.03903	0.39211	0.03938	0.03926
23	0.10886	0.10915	0.10967	0.11017	0.11067	0.11034
22	0.20551	0.20464	0.20418	0.20372	0.20327	0.20357
33	0.11830	0.11779	0.11753	0.11727	0.11701	0.11718

Table 3
Values of $T_n^{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ for different values of λ_1 and λ_2.

	$\lambda_2 = -1/2$	$\lambda_2 = 0$	$\lambda_2 = 1/2$	$\lambda_2 = 1$	$\lambda_2 = 2/3$
$\lambda_1 = -1$	17.14	17.41	22.193	24.588	23.144
$\lambda_1 = -1/2$	32.77	31.899	31.722	31.669	31.699
$\lambda_1 = 0$	31.84	31.627	31.541	31.53	31.532
$\lambda_1 = 2/3$	32.77	31.379	31.722	31.669	31.699
$\lambda_1 = 1$	31.66	31.452	31.297	31.129	31.235

i.e., the minimum discrimination estimator for symmetry introduced and studied in [2].

For $\lambda = 1$,

$$\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}^{(1)}) = \frac{\left(\frac{\widehat{p}_{ij}^2+\widehat{p}_{ji}^2}{2}\right)^{1/2}}{\sum_{i=1}^{3}\sum_{j=1}^{3}\left(\frac{\widehat{p}_{ij}^2+\widehat{p}_{ji}^2}{2}\right)^{1/2}},$$

and we get the minimum chi-squared estimator for symmetry introduced in [3].

Other interesting estimators for symmetry are as follows. For $\lambda_2 = -2$, the minimum modified chi-squared estimator; for $\lambda = -1/2$, the minimum Freeman–Tukey estimator; and finally, for $\lambda_2 = 2/3$, the minimum Cressie and Read estimator.

In Table 2, we present the expression of $\boldsymbol{p}_{ij}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ for different values of λ_2 ($\lambda_2 = -1, -1/2, 0, 1/2, 1, 2/3$).

In relation with (27), we get for $\lambda_1 = \lambda_2 = 0$ the likelihood ratio test,

$$T_n^{(0)}(\widehat{\boldsymbol{\theta}}^{(0)}) = G^2 = 2\sum_{\substack{i,j\\ i<j}} n_{ij}\log\frac{2n_{ij}}{n_{ji}+n_{ij}}.$$

For $\lambda_2 = 0$ and $\lambda_1 = 1$, we get the classical chi-square test statistic given in [1]

$$T_n^{(1)}(\widehat{\boldsymbol{\theta}}^{(0)}) = X^2 = \sum_{\substack{i,j\\ i<j}} \frac{\left(n_{ij}-n_{ji}\right)^2}{n_{ij}+n_{ji}}.$$

Finally, for $\lambda_1 = \lambda_2 = 1$, we get the test statistics given in [3] and for $\lambda_2 = 0$ we get the family of test statistics introduced and studied in [17].

Table 3 presents the values of different test statistics for different values of λ_1 and λ_2.

Based on the results given in Table 3, we must reject the hypothesis of symmetry in our data, and therefore it will be not good to consider the restricted estimators $\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{(\lambda_1)})$. If we observe the expression of the preliminary test estimator given in (9), we can observe that $\boldsymbol{p}_{(\lambda_1)}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ coincides with the unrestricted estimator $\widehat{\boldsymbol{p}} = \left(\frac{n_{11}}{n}, \ldots, \frac{n_{33}}{n}\right)$, i.e.,

$$\boldsymbol{p}_{(\lambda_1)}^{\mathrm{pte}}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) = (0.16814, 0.11295, 0.04568, 0.0980, 0.20418, 0.12782, 0.03259, 0.09150, 0.11753).$$

In order to get the expression of the shrinkage preliminary phi-divergence estimator, $\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$, we need to fix a value of $a \in (0, 1)$. We consider $a = 0.4$. Then,

$$\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) = 0.6\boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) + 0.4\widehat{\boldsymbol{p}}.$$

In Table 4, we present the expression of $\boldsymbol{p}^{\mathrm{sre}}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ for different values of λ_2.

In Table 5, we present the expressions of the James–Stein phi-divergence estimator, $\boldsymbol{p}_{(\lambda_1)}^{s}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$. In order to get $\boldsymbol{p}_{(\lambda_1)}^{s}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$, we need $T_n^{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$. We have used the values of $T_n^{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ for $\lambda_1 = 1$ and $\lambda_2 = -1/2$. In our situation,

Table 4
Srinkage preliminary phi-divergence estimators for different values of λ_2 $(-1, -1/2, 0, 1/2, 1, 2/3)$.

Cell	$\lambda_2 = -1$	$\lambda_2 = -1/2$	$\lambda_2 = 0$	$\lambda_2 = 1/2$	$\lambda_2 = 1$	$\lambda_2 = 2/3$
11	0.16858	0.16829	0.16814	0.16794	0.16784	0.16794
12	0.11052	0.11037	0.11032	0.11027	0.11021	0.11025
13	0.04427	0.04282	0.04289	0.18412	0.04303	0.04298
21	0.10263	0.10248	0.10243	0.10238	0.10232	0.10236
22	0.20466	0.20432	0.20413	0.20395	0.20377	0.20389
23	0.12024	0.12035	0.12056	0.12076	0.12096	0.12083
31	0.03505	0.03509	0.03517	0.17640	0.03531	0.03526
32	0.09844	0.09856	0.09877	0.09897	0.11047	0.09904
33	0.11784	0.11763	0.11753	0.11743	0.11732	0.11739

Table 5
James–Stein preliminary phi-divergence estimators for different values of λ_2 $(-1, -1/2, 0, 1/2, 1, 2/3)$.

Cell	$\lambda_2 = 1$	$\lambda_2 = -1/2$	$\lambda_2 = 0$	$\lambda_2 = 1/2$	$\lambda_2 = 1$	$\lambda_2 = 2/3$
11	0.1686	0.16858	0.16857	0.16855	0.16854	0.16855
12	0.1104	0.11039	0.11039	0.11038	0.11038	0.11038
13	0.04409	0.04409	0.04410	0.06227	0.04411	0.04411
21	0.10276	0.10275	0.10275	0.10274	0.10274	0.10274
22	0.20469	0.20466	0.20464	0.20463	0.20462	0.20463
23	0.12018	0.11989	0.11991	0.11992	0.11994	0.11993
31	0.03517	0.03517	0.03518	0.04633	0.03.51	0.03743
32	0.09877	0.09878	0.09880	0.09881	0.09883	0.09882
33	0.11785	0.11784	0.11783	0.11782	0.11781	0.11782

$T_n^{(1)}(\widehat{\boldsymbol{\theta}}^{(-1/2)}) = 31.66.$

$$\boldsymbol{p}^{s}_{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) = \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) + \left(1 - T_n^{(1)}(\widehat{\boldsymbol{\theta}}^{(-1/2)})\right)\left(\widehat{\boldsymbol{p}} - \boldsymbol{p}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})\right)$$

using λ_1 and λ_2, i.e., $T_n^{(1)}(\widehat{\boldsymbol{\theta}}^{(-1/2)})$.

In our case, $m = 3$. Therefore,

$$I_{(m-2,\infty)}\left(T_n^{(1)}(\widehat{\boldsymbol{\theta}}^{(-1/2)})\right) = 1$$

and $\boldsymbol{p}^{s+}_{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$ coincides with $\boldsymbol{p}^{s}_{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$. We also have

$$I_{(m-2,\infty)}(x) = I_{\left[\chi^2_3,\infty\right)}(x),$$

and therefore $\boldsymbol{p}^{\text{pte}_+}_{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)}) = \boldsymbol{p}^{s+}_{(\lambda_1)}(\widehat{\boldsymbol{\theta}}^{(\lambda_2)})$.

Acknowledgement

We would like to thank an anonymous referee for helpful comments since his/her interesting suggestions improved the paper. This work was supported by Grant MTM2009-10072.

References

[1] A. Bowker, A test for symmetry in contingency tables, Journal of the American Statistical Association 43 (1948) 572–574.
[2] C.T. Ireland, H.H. Ku, S. Kullback, Symmetry and marginal homogeneity of an $r \times r$ contingency table, Journal of the American Statistical Association 64 (1969) 1323–1341.
[3] D. Quade, I.A. Salama, A note on minimum chi-square statistics in contingency tables, Biometrics 31 (1975) 953–956.
[4] T.R.C. Read, N.A.C. Cressie, Goodness of Fit Statistics for Discrete Multivariate Data, Springer-Verlag, New York, 1988.
[5] 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, Journal of Statistical Computation and Simulation 75 (7) (2005) 555–580.
[6] T.A. Bancroft, On biases in estimation due to the use of preliminary tests of significance, Annals of Mathematical Statistics 15 (1944) 190–204.
[7] C.P. Han, T.A. Bancroft, On pooling means when variance is unknown, Journal of the American Statistical Association 63 (1986) 1333–1342.
[8] C. Stein, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, University of California Press, Berkeley, CA, 1956, pp. 197–206.
[9] A.K.Md.E. Saleh, Theory of Preliminary Test and Stein-Type Estimation with Applications, Wiley, 2006.
[10] D. Morales, L. Pardo, I. Vajda, Asymptotic divergences of estimates of discrete distributions, Journal of Statistical Planning and Inference 48 (1995) 347–369.
[11] S.M. Ali, S.D. Silvey, A general class of coefficients of divergence of one distribution from another, Journal of the Royal Statistical Society, Series B 26 (1966) 131–142.
[12] I. Csiszàr, Eine informationstheorestiche Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, Publications of the Mathematical Institute of Hungarian Academy of Sciences, Series A 8 (1963) 84–108.

[13] L. Pardo, Statistical Inference Based on Divergence Measures, Chapman & Hall, CRC, New York, 2006.
[14] M.L. Menéndez, L. Pardo, M.C. Pardo, Preliminary test estimators and phi-divergence measures in generalized linear models with binary data, Journal of Multivariate Analysis 99 (2008) 2265–2284.
[15] L. Pardo, M.L. Menéndez, On some pre-test and Stein-rule phi-divergence test estimators in the independence model of categorical data, Journal of Statistical Planning and Inference 138 (2007) 2163–2179.
[16] D.V. Glass, Social Mobility in Britain, Free Press, Glencoe, Illinois, 1954.
[17] M.L. Menéndez, J.A. Pardo, L. Pardo, Tests based on phi-divergences for bivariate symmetry, Metrika 53 (2001) 15–29.