Kybernetika

Teresa Pérez; Julio A. Pardo

Goodness-of-fit tests based on K_ϕ-divergence

Kybernetika, Vol. 39 (2003), No. 6, [739]--752

Persistent URL: http://dml.cz/dmlcz/135569

KYBERNETIKA — VOLUME 39 (2003), NUMBER 6, PAGES 739–752

GOODNESS-OF-FIT TESTS BASED ON K_ϕ-DIVERGENCE

TERESA PÉREZ AND JULIO A. PARDO

In this paper a new family of statistics based on K_ϕ-divergence for testing goodness-of-fit under composite null hypotheses are considered. The asymptotic distribution of this test is obtained when the unspecified parameters are estimated by maximum likelihood as well as minimum K_ϕ-divergence.

Keywords: K_ϕ-divergence, goodness-of-fit, minimum K_ϕ-divergence estimate
AMS Subject Classification: 62B10, 62E20

1. INTRODUCTION

It is known that the problem of testing goodness-of-fit of statistical data with the hypothesis that the law of distribution of the observable random variable belongs to a given class of distributions, can be reduced by means of grouping observations to the analogous problem for a discrete distribution. This is possible when we group the model into m classes $C_1, C_2, ..., C_m$ with corresponding probabilities $\pi_1, \pi_2, ..., \pi_m$. In this case, the general goodness-of-fit problem reduces to testing a hypothesis about the parameter $\pi = (\pi_1, \pi_2, ..., \pi_m)^t$ from a multinomial random variable $X = (X_1, X_2, ..., X_m)^t$ of parameters $(n, \pi_1, \pi_2, ..., \pi_m)^t$, i.e.,

$$H_0 : \pi = \pi_0 = (\pi_{01}, ..., \pi_{0m})^t \in \Pi_0 \tag{1.1}$$

where $\Pi_0 \subset \Delta_m = \{P = (p_1, ..., p_m),\ p_i > 0 \text{ and } \sum_{i=1}^m p_i = 1\}$ is the null model space of probability vectors. The null hypothesis may completely specify π, e.g., a simple hypothesis. Otherwise the null hypothesis is composite, specifying π as a function of a smaller number of unknown parameters which needs to be estimated from the experimental data.

To solve these problems Cressie and Read [5] and Read and Cressie [15] proposed a generalized statistic which they called the power divergence statistic. This family contains the Pearson's chi-square statistic, the loglikelihood ratio statistic, the Freeman Tukey statistic, and the modified likelihood ratio. Zografos et al [17] proposed to use a family of statistics based on the Csiszár divergence family [6] to solve this problem under simple hypotheses and Morales et al [9] under composite hypothesis. This family contains the power divergence statistic for

$\varphi(x) = \varphi_\lambda(x) = (\lambda(\lambda+1))^{-1}\left(x^{\lambda+1} - (\lambda+1)x + \lambda\right)$, $\lambda \in R$, $\lambda \neq -1, 0$ and it is given for two probability distributions π_1 and π_2 by

$$D^{\varphi}(\pi_1, \pi_2) = \sum_{i=1}^{m} \pi_{2i} \varphi\left(\frac{\pi_{1i}}{\pi_{2i}}\right),$$

for a convex function, $\varphi : [0, \infty) \to \mathbb{R} \cup \{\infty\}$, where $0\varphi(0/0) = 0$ and $0\varphi(p/0) = \lim_{n\to\infty}(\varphi(u)/u)$.

M. C. Pardo [10] consider for this proposal a family of statistic based on the R_ϕ divergence introduced and studied by Rao [13], Burbea and Rao [4] and Burbea [3]. This family is defined, for two probability distributions π_1 and π_2, as

$$R_\phi(\pi_1, \pi_2) = H_\phi\left(\frac{\pi_1 + \pi_2}{2}\right) - \frac{1}{2}\left(H_\phi(\pi_1) + H_\phi(\pi_2)\right)$$

where $H_\phi(\pi)$ is the ϕ-entropy introduced by Burbea and Rao [4] so that $H_\phi(\pi) = -\sum_{i=1}^{m} \phi(\pi_i)$ with $\phi : (0, \infty) \to R$ being convex function. Some interesting properties of the ϕ-entropies can be seen in Vajda and Vašek [16].

If any case we can observe that implicitly or explicitly the goodness-of-fit tests are based on distances, dissimilarities, or simply divergences. For this reason, we can use measures of divergences different from the φ-divergences and the R_ϕ-divergences.

There is an important family of divergences, the K_ϕ-divergence, introduced and studied by Burbea and Rao [4]. This family is define for two probability distributions π_1 and π_2 as

$$K_\phi(\pi_1, \pi_2) = \sum_{i=1}^{m} (\pi_{1i} - \pi_{2i}) \left[\frac{\phi(\pi_{1i})}{\pi_{1i}} - \frac{\phi(\pi_{2i})}{\pi_{2i}}\right]$$

where ϕ is a convex function defined in an interval I not containing the origin. The convexity of the K_ϕ-divergence is obtained if the following property holds

$$[\phi''(p_i) - q_i\varphi''(p_i)] \times [\phi''(q_i) - p_i\varphi''(q_i)] - [\varphi'(p_i) + \varphi'(q_i)]^2 \geq 0, \;\; i = 1, 2, ..., m.$$

An important family of K_ϕ-divergences, studied in Burbea and Rao [4], is obtained if we consider the function

$$\phi(x) = \phi_\tau(x) = \begin{cases} x \log x & \text{if } \tau = 1 \\ (\tau - 1)^{-1}(x^\tau - x) & \text{if } \tau \in (1, 2] \end{cases}$$

In this case the K_{ϕ_τ}-divergence is convex if $\tau \in [1, 2]$.

On the basis of the K_ϕ-divergence between the observed proportions X/n and the hypothec proportions π_0, one can introduce a statistic for the goodness-of-fit problem which will be denoted by $K_\phi(X/n, \pi_0)$. Under the simple hypothesis, Pérez and Pardo [11] established, for $\pi_0 = (1/m, ..., 1/m)^t$ that $\dfrac{mnK_\phi(X/n, \pi_0)}{\varphi'(1/m)}$ with $\varphi(x) = \phi(x)/x$ is asymptotically chi-square distributed with $m - 1$ degrees of freedom.

In the next section we obtain the asymptotic distribution formula of the statistic $nK_\phi(X/n, \hat{\pi})$ under composite hypotheses. Here it is possible to estimate π by the maximum likelihood method but we propose to use the minimum K_ϕ-divergence estimate, analyzed by Pérez and Pardo [12], defined as $\hat{\pi} = \arg\min K_\phi(X/n, \pi)$.

2. ASYMPTOTIC DISTRIBUTION OF K_ϕ-DIVERGENCE STATISTICS

Let $X = (X_1, ..., X_m)^t$ be a random vector with multinomial distribution $M_m(n, \pi)$ and consider the null hypotheses

$$H_0 : \pi = f(\vartheta) \in \Pi_0 \subset \Delta_m \tag{2.1}$$

where $\Pi_0 = \{f(\vartheta) : \vartheta \in \Theta_0\}$ with $f(\vartheta) = (f_1(\vartheta), ..., f_m(\vartheta))^t$ and $\vartheta = (\vartheta_1, \vartheta_2 ..., \vartheta_s)^t \in \Theta_0 \subset R^s$ is an unknown parameters vector. To solve this goodness of fit problem it is necessary to choose an estimate $\hat{\pi} \in \Pi_0$ so that it is as close as possible to X/n and we assume that the model is correct, i.e., exist a value $\vartheta^* \in \Theta_0$ with $f(\vartheta^*) = \pi^* \in \Pi_0$, where π^* is the true value of the multinomial probability. The most commonly way of finding the value of $\hat{\pi}$ is the maximum likelihood method, that it is equivalent to minimize over $\vartheta \in \Theta_0 \subset R^s$ the Kullback divergence given by

$$D(X/n, f(\vartheta)) = \sum_{i=1}^{m} X_i/n \log \frac{X_i/n}{f_i(\vartheta)}.$$

In general, we can choose as estimator of ϑ the value $\hat{\vartheta}$ verifying

$$D(X/n, f(\hat{\vartheta})) = \inf_{\vartheta \in \Theta_0} D(X/n, f(\vartheta))$$

where D is a given divergence measure. Depending on the chosen divergence measure, different estimators are obtained. On one hand, if

$$D(\hat{P}, f(\vartheta)) = n \sum_{i=1}^{m} \frac{(\hat{p}_i - f_i(\vartheta))^2}{f_i(\vartheta)}$$

then the corresponding $\hat{\vartheta}$ is the well-known minimum χ^2 estimator, studied in this context by Fryer and Robertson [8]. On the other hand, if we consider the K_ϕ-divergence, the corresponding $\hat{\vartheta}$ will be called the minimum K_ϕ-divergence estimator.

The following definition was given in Pérez and Pardo [12].

Definition 2.1. The minimum K_ϕ-divergence estimator of ϑ^* is any $\hat{\vartheta}_\phi \in \bar{\Theta}_0$ (the closure of Θ_0) such that $\hat{\pi}_\phi = f(\hat{\vartheta}_\phi)$ and

$$K_\phi(X/n, f(\hat{\vartheta}_\phi)) = \inf_{\vartheta \in \Theta_0} K_\phi(X/n, f(\vartheta)).$$

So the minimum K_ϕ-divergence estimator is $\hat{\vartheta}_\phi = \arg\min K_\phi(X/n, f(\vartheta))$.

As it can be seen in Pérez and Pardo [12] in order to assure that $\hat{\vartheta}_\phi$ exists and is consistent for ϑ^* it suffices, under (2.1) to satisfy regularity conditions of Birch [1]. Moreover, if φ is a convex function and $\varphi(x) = \phi(x)/x$ is concave, we have the following properties

(1)
$$\hat{\vartheta}_\phi = \vartheta^* + (A_D^t A_D)^{-1} A_D^t \left(D_{\varphi'(\pi^*)}\right)^{\frac{1}{2}} (X/n - \pi^*) + o_p(n^{\frac{-1}{2}}) \quad (2.2)$$

(2) The asymptotic distribution of $\sqrt{n}\left(\hat{\vartheta}_\phi - \vartheta^*\right)$ is

$$N\Big(0, (A_D A_D)^{-1} A_D^t \left(D_{\varphi'(\pi^*)}\right)^{-\frac{1}{2}} \left(D_{\pi^*} - \pi^*\pi^{*t}\right) \left(D_{\varphi'(\pi^*)}\right)^{-\frac{1}{2}} A_D (A_D^t A_D)^{-1}\Big) \quad (2.3)$$

with $A_D = (A_{Dij})$, $A_{Dij} = \left(\frac{\partial f_i(\vartheta^*)}{\partial \vartheta_j}\right) \varphi'(\pi_i^*)^{\frac{1}{2}} \quad i = 1,..,m;\ j = 1,...,s,$
where $D_{(c_1,\dots,c_m)} = \operatorname{diag}(c_1, ..., c_m)$ for any $(c_1, ..., c_m) \in \mathbb{R}^m$.

If we want to use the statistic $K_\phi(X/n, \hat{\pi}_\phi))$ for testing (2.1) we need to know its asymptotic distribution under the null hypotheses so that we can construct a critical region.

Theorem 2.1. If ϕ is a convex function, $\varphi(x) = \phi(x)x$ is concave and $\hat{\pi}_\phi \in \Pi_0$ is the minimum K_ϕ-divergence estimator of $\pi^* = f(\vartheta^*)$ then, under the regularity conditions of Birch [1], $W_n^* = \sqrt{n}\,(X/n - \hat{\pi}_\phi)$ converges in distribution to a multivariate normal random vector W^* as $n \to \infty$, with mean vector 0 and covariance matrix

$$\Sigma_D(\vartheta^*) =$$
$$= \left(D_{\pi^*} - \pi^*\pi^{*t} - \left(D_{\pi^*} - \pi^*\pi^{*t}\right) L^t - L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) + L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) L^t\right) \quad (2.4)$$

where

$$L = \left(D_{\varphi'(\pi^*)}\right)^{-\frac{1}{2}} A_D (A_D^t A_D)^{-1} A_D^t \left(D_{\varphi'(\pi^*)}\right)^{-\frac{1}{2}} \quad (2.5)$$

with

$$A_{Dij} = \left(\frac{\partial f_i(\vartheta^*)}{\partial \vartheta_j}\right) \varphi'(\pi_i^*)^{\frac{1}{2}}.$$

Proof. From Birch's conditions we get

$$f(\hat{\vartheta}_\phi) - f(\vartheta^*) = \left(\frac{\partial f(\vartheta^*)}{\partial \vartheta}\right) \left(\hat{\vartheta}_\phi - \vartheta^*\right) + o_p(n^{-\frac{1}{2}})$$

since $\left\|\hat{\vartheta}_\phi - \vartheta^*\right\| = O_p(n^{-\frac{1}{2}})$ by (2.3). Consequently by (2.2)

$$f(\hat{\vartheta}_\phi) - f(\vartheta^*) = \left(\frac{\partial f(\vartheta^*)}{\partial \vartheta}\right)(A_D^t A_D)^{-1} A_D^t \left(D_{\varphi'(\pi^*)}\right)^{\frac{1}{2}} (X/n - \pi^*) + o_p(n^{-\frac{1}{2}}) =$$

$$= \left(D_{\varphi'(\pi^*)}\right)^{-\frac{1}{2}} A_D (A_D^t A_D)^{-1} A_D^t \left(D_{\varphi'(\pi^*)}\right)^{\frac{1}{2}} (X/n - \pi^*) + o_p(n^{-\frac{1}{2}}).$$

We can write

$$\begin{pmatrix} X/n - \pi^* \\ \hat{\pi}_\phi - \pi^* \end{pmatrix} = \begin{pmatrix} I \\ L \end{pmatrix} (X/n - \pi^*) + o_p(n^{-\frac{1}{2}}) \tag{2.6}$$

with L defined in (2.5) and I the identity $m \times m$ matrix. As $\sqrt{n}\,(X/n - \pi^*)$ has an asymptotic normal distribution with mean zero and covariance matrix $D_{\pi^*} - \pi^* \pi^{*t}$, we get

$$\sqrt{n}\begin{pmatrix} X/n - \pi^* \\ \hat{\pi}_\phi - \pi^* \end{pmatrix} \xrightarrow[n\to\infty]{L} N\left(0, \begin{pmatrix} D_{\pi^*} - \pi^*\pi^{*t} & (D_{\pi^*} - \pi^*\pi^{*t})\,L^t \\ L\,(D_{\pi^*} - \pi^*\pi^{*t}) & L\,(D_{\pi^*} - \pi^*\pi^{*t})\,L^t \end{pmatrix}\right)$$

Using multinomial normal distributions properties we get that the asymptotic distribution of the vector

$$\sqrt{n}\,(X/n - \pi^*) - \sqrt{n}\,(\hat{\pi}_\phi - \pi^*) = \sqrt{n}\,(X/n - \hat{\pi}_\phi)$$

is normal with mean zero and covariance matrix $\Sigma_D(\vartheta^*)$ defined in (2.4).

Theorem 2.2. If ϕ_A, ϕ_B are convex functions and $\varphi_A(x) = \phi_A(x)/x$, $\varphi_B(x) = \phi_B(x)/x$ are concave, then

$$nK_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) \xrightarrow[n\to\infty]{L} \sum_{i=1}^{r} \xi_i(\vartheta^*) Z_i^2$$

where $\xi_i(\vartheta^*)$ are the nonzero eigenvalues of $D_{\varphi'_A(\pi^*)}\Sigma_D(\vartheta^*)$, with $\Sigma_D(\vartheta^*)$ defined in (2.4), $r = \operatorname{rank}(\Sigma_D(\vartheta^*) D_{\varphi'_A(\pi^*)} \Sigma_D(\vartheta^*))$ and Z_i are independent standard normal variables.

Proof. Define

$$F(x) = (x - \hat{\pi}_{i\phi_B})\left[\frac{\phi(x)}{x} - \frac{\phi(\hat{\pi}_{i\phi_B})}{\hat{\pi}_{i\phi_B}}\right].$$

Taylor series expansion of $F(X_i/n)$ around $\hat{\pi}_{i\phi_B}$ yields

$$F(X_i/n) = F(\hat{\pi}_{i\phi_B}) + F'(\hat{\pi}_{i\phi_B})(X_i/n - \hat{\pi}_{i\phi_B}) + F''(\hat{\pi}_{i\phi_B})\frac{(X_i/n - \hat{\pi}_{i\phi_B})^2}{2} + o_p(n^{-1}) =$$

$$= \phi'(\hat{\pi}_{i\phi_B})(X_i/n - \hat{\pi}_{i\phi_B})^2 + o_p(n^{-1}),$$

so that

$$K_\phi(X/n, \hat{\pi}_{\phi_B}) = \sum_{i=1}^{m}(X_i/n - \hat{\pi}_{i\phi_B})^2\varphi'(\hat{\pi}_{i\phi_B}) + o_p(n^{-1}).$$

Therefore

$$nK_\phi(X/n, \hat{\pi}_{\phi_B}) = W_n^{*t} D_{\varphi'_A(\hat{\pi}_{\phi_B})} W_n^* + o_p(1)$$

where $W_n^* = \sqrt{n}(X/n - \hat{\pi}_{\phi_B})$ and $D_{\varphi'_A(\hat{\pi}_{\phi_B})} = diag\left(\varphi'_A(\hat{\pi}_{1\phi_B}), ..., \varphi'_A(\hat{\pi}_{m\phi_B})\right)$.

Since $\hat{\pi}_{\phi_B}$ is the minimum K_ϕ-divergence estimate, it is consistent in the sense $\hat{\pi}_{\phi_B} = \pi^* + o_p(1)$ (see Pérez and Pardo [12]). This implies that

$$nK_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) = W_n^{*t} D_{\varphi'_A(\hat{\pi}_{\phi_B})} W_n^* + o_p(1) \xrightarrow[n\to\infty]{L} W^{*t} D_{\varphi'_A(\pi^*)} W^*$$

where $W^* = \sqrt{n}\,(X/n - \pi^*)$.

From Theorem 2.1, we know that W^* has the multivariate normal distribution with mean vector 0 and covariance matrix $\Sigma_D(\vartheta^*)$ defined in [2.4]. So the distribution of $W^{*t} D_{\varphi'_A(\pi^*)} W^*$ is $\sum_{i=1}^{r} \xi_i(\vartheta^*) Z_i^2$ where $\xi_i(\vartheta^*)$ are the nonzero eigenvalues of $D_{\varphi'_A(\pi^*)}\Sigma_D(\vartheta^*)$, $r = \text{rank}\,(\Sigma_D(\vartheta^*) D_{\varphi'_A(\pi^*)} \Sigma_D(\vartheta^*))$ and Z_i are independent standard normal variables. □

Corollary 2.1. If $\phi(x) = \phi_1(x) = x \log x$, then

$$nK_{\phi_1}(X/n, \hat{\pi}_{\phi_1}) \xrightarrow[n\to\infty]{L} \chi^2_{m-s-1}.$$

Proof. In Theorem 2.2 we have proved that

$$nK_{\phi_1}(X/n, \hat{\pi}_{\phi_1}) \xrightarrow[n\to\infty]{L} \sum_{i=1}^{r} \xi_i(\vartheta^*) Z_i^2$$

where $r = \text{rank}\,(\Sigma_D(\vartheta^*) D_{(\pi^*)^{-1}} \Sigma_D(\vartheta^*))$ with $\Sigma_D(\vartheta^*)$ defined as in (2.4), Z_i are independent standard normal variables and $\xi_i(\vartheta^*)$ are the nonzero eigenvalues of $T = D_{(\pi^*)^{-1}}\left(D_{\pi^*} - \pi^*\pi^{*t} - (D_{\pi^*} - \pi^*\pi^{*t})\,L^t - L\,(D_{\pi^*} - \pi^*\pi^{*t}) + L\,(D_{\pi^*} - \pi^*\pi^{*t})\,L^t\right)$.

The eigenvalues of T are the same as the eigenvalues of

$$\begin{aligned} T^* &= D_{(\pi^*)^{-\frac{1}{2}}} D_{\pi^*} - \pi^*\pi^{*t} - \left(D_{\pi^*} - \pi^*\pi^{*t}\right) L^t - \\ &- L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) + L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) L^t D_{(\pi^*)^{-\frac{1}{2}}} = \\ &= D_{(\pi^*)^{-\frac{1}{2}}}(D_{\pi^*} - \pi^*\pi^{*t}) D_{(\pi^*)^{-\frac{1}{2}}} - D_{(\pi^*)^{-\frac{1}{2}}}\left(D_{\pi^*} - \pi^*\pi^{*t}\right) L^t D_{(\pi^*)^{-\frac{1}{2}}} - \\ &- D_{(\pi^*)^{-\frac{1}{2}}} L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) D_{(\pi^*)^{-\frac{1}{2}}} + D_{(\pi^*)^{-\frac{1}{2}}} L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) L^t D_{(\pi^*)^{-\frac{1}{2}}}. \end{aligned}$$

Since $\pi^{*t} = \sqrt{\pi^{*t}} D_{(\pi^*)^{\frac{1}{2}}}$, we have

$$\begin{aligned} S &= D_{(\pi^*)^{-\frac{1}{2}}}\left(D_{\pi^*} - \pi^*\pi^{*t}\right) D_{(\pi^*)^{-\frac{1}{2}}} = \\ &= D_{(\pi^*)^{-\frac{1}{2}}} D_{\pi^*} D_{(\pi^*)^{-\frac{1}{2}}} - D_{(\pi^*)^{-\frac{1}{2}}} \pi^*\pi^{*t} D_{(\pi^*)^{-\frac{1}{2}}} = I - \sqrt{\pi^*}\sqrt{\pi^{*t}}. \end{aligned}$$

and

$$\begin{aligned} B &= D_{(\pi^*)^{-\frac{1}{2}}} L \left(D_{\pi^*} - \pi^*\pi^{*t}\right) D_{(\pi^*)^{-\frac{1}{2}}} = \\ &= D_{(\pi^*)^{-\frac{1}{2}}} L D_{(\pi^*)^{\frac{1}{2}}} - D_{(\pi^*)^{-\frac{1}{2}}} L(\pi^*\pi^{*t}) D_{(\pi^*)^{-\frac{1}{2}}} = \\ &= D_{(\pi^*)^{-\frac{1}{2}}} L D_{(\pi^*)^{\frac{1}{2}}} (I - \sqrt{\pi^*}\sqrt{\pi^{*t}}) = SK \end{aligned}$$

with

$$\begin{aligned} K &= D_{(\pi^*)^{-\frac{1}{2}}} L D_{(\pi^*)^{\frac{1}{2}}} = \\ &= D_{(\pi^*)^{-\frac{1}{2}}} \left(D_{(\pi^*)^{1/2}}\right) A_D (A_D^t A_D)^{-1} A_D^t \left(D_{(\pi^*)^{1/2}}\right) D_{(\pi^*)^{\frac{1}{2}}} = \\ &= A_D (A_D^t A_D)^{-1} A_D^t. \end{aligned}$$

Hence

$$\begin{aligned} T^* &= S - A_D (A_D^t A_D)^{-1} A_D^t S - \\ &- S A_D (A_D^t A_D)^{-1} A_D^t + A_D (A_D^t A_D)^{-1} A_D^t S A_D (A_D^t A_D)^{-1} A_D^t = \\ &= (I - \sqrt{\pi^*}\sqrt{\pi^{*t}}) - A_D (A_D^t A_D)^{-1} A_D^t (I - \sqrt{\pi^*}\sqrt{\pi^{*t}}) - \\ &- (I - \sqrt{\pi^*}\sqrt{\pi^{*t}}) A_D (A_D^t A_D)^{-1} A_D^t + \\ &+ A_D (A_D^t A_D)^{-1} A_D^t (I - \sqrt{\pi^*}\sqrt{\pi^{*t}}) A_D (A_D^t A_D)^{-1} A_D^t. \end{aligned}$$

As it is verified that $\sqrt{\pi^{*t}} A_D = 0$ we get

$$T^* = I - \sqrt{\pi^*}\sqrt{\pi^{*t}} - A_D (A_D^t A_D)^{-1} A_D^t.$$

Now we are going to prove that T^* is idempotent

$$\begin{aligned} (T^*)^2 &= I - \sqrt{\pi^*}\sqrt{\pi^{*t}} - A_D (A_D^t A_D)^{-1} A_D^t - \\ &- \sqrt{\pi^*}\sqrt{\pi^{*t}} + \sqrt{\pi^*}\sqrt{\pi^{*t}}\sqrt{\pi^*}\sqrt{\pi^{*t}} + \sqrt{\pi^*}\sqrt{\pi^{*t}} A_D (A_D^t A_D)^{-1} A_D^t - \\ &- A_D (A_D^t A_D)^{-1} A_D^t + A_D (A_D^t A_D)^{-1} A_D^t \sqrt{\pi^*}\sqrt{\pi^{*t}} + \\ &+ A_D (A_D^t A_D)^{-1} A_D^t A_D (A_D^t A_D)^{-1} A_D^t. \end{aligned}$$

As $A_D (A_D^t A_D)^{-1} A_D^t \sqrt{\pi^*}\sqrt{\pi^{*t}} = 0$ and $\sqrt{\pi^{*t}}\sqrt{\pi^*} = 1$, we get

$$(T^*)^2 = I - \sqrt{\pi^*}\sqrt{\pi^{*t}} - A_D (A_D^t A_D)^{-1} A_D^t = T^*.$$

Since T^* is idempotent, its eigenvalues are either 0 or 1. The number of nonzero eigenvalues is equal to

$$\begin{aligned} \text{Trace}(T^*) &= \text{Trace}\,(I) - \text{Trace}\,(\sqrt{\pi^*}\sqrt{\pi^{*t}}) - \text{Trace}\,(A_D (A_D^t A_D)^{-1} A_D^t) = \\ &= m - 1 - \text{Trace}\,(A_D (A_D^t A_D)^{-1} A_D^t) = m - 1 - s, \end{aligned}$$

since $(A_D^t A_D)^{-1} A_D^t A_D = I_{s\times s}$. □

Now we consider the maximum likelihood estimate, $\tilde{\pi}$ instead of the minimum K_ϕ-divergence. In the next theorem we calculate the asymptotic distribution of $nK_\phi(X/n, \tilde{\pi})$.

Theorem 2.3. a) Under Birch's regularity conditions, if $\tilde{\pi} \in \Pi_0$ is the maximum likelihood estimate of $\pi^* = f(\vartheta^*)$, then

$$W_n^* = \sqrt{n}(X/n - \tilde{\pi})$$

converges in distribution to a multivariate normal vector W^* as $n \to \infty$, with mean vector 0 and covariance matrix

$$\Sigma_L(\vartheta^*) = D_{\pi^*} - \pi^*\pi^{*t} - \left(D_{(\pi^*)^{1/2}}\right) A_L(A_L^t A_L)^{-1} A_L^t \left(D_{(\pi^*)^{1/2}}\right) \qquad (2.7)$$

where A_L is an $m \times s$ matrix so that $(A_L)_{ij} = \left(\dfrac{\partial f_i(\vartheta^*)}{\partial \vartheta_j}\right)(\pi_i^*)^{-\frac{1}{2}}$.

b) Let ϕ be a convex function and $\varphi(x) = \phi(x)/x$ concave. Then

$$nK_\phi(X/n, \tilde{\pi}) \xrightarrow[n\to\infty]{L} \sum_{i=1}^{r} \xi_i(\theta^*) Z_i^2$$

where $r = \operatorname{rank}(\Sigma_L(\vartheta^*) D_{\varphi'(\pi^*)} \Sigma_L(\vartheta^*))$ with $\Sigma_L(\vartheta^*)$ defined as in (2.7), Z_i are independent standard normal variables and $\xi_i$are the nonzero eigenvalues of

$$D_{\varphi'(\pi^*)}\left(D_{\pi^*} - \pi^*\pi^{*t} - \left(D_{(\pi^*)^{1/2}}\right) A_L(A_L^t A_L)^{-1} A_L^t \left(D_{(\pi^*)^{1/2}}\right)\right)$$

c) For $\phi = \phi_\tau$ with $\tau = 1$, the statistic

$$nK_{\phi_1}(X/n, \tilde{\pi}) \xrightarrow[n\to\infty]{L} \chi^2_{m-s-1}.$$

Proof. Using the asymptotic normality of the BAN estimate $\tilde{\pi}$, it is straightforward to parallel the results given for the minimum K_ϕ-divergence estimate.

a) See Theorem 14.8-4 of Bishop et al [2], pp. 511.

b) In the same way as in Theorem 2.2 we write

$$nK_\phi(X/n, \tilde{\pi}) = \sum_{i=1}^{m} n(X_i/n - \tilde{\pi}_i)^2 \varphi'(\tilde{\pi}_i) = W_n^{*t} D_{\varphi'(\tilde{\pi})} W_n^* + o_p(1) \xrightarrow[n\to\infty]{L} W^{*t} D_{\varphi'(\pi^*)} W^*$$

since $\tilde{\pi} = \pi^* + o_p(1)$ because the maximum likelihood estimate is BAN. By a) we know that the asymptotic distribution of W^* is normal multivariate with mean vector 0 and covariance matrix $\Sigma_L(\vartheta^*)$.

So the distribution of $W^{*t} D_{\varphi'(\pi^*)} W^*$ is equal to $\sum_{i=1}^{r} \xi_i(\vartheta^*) Z_i^2$ where r is given by $r = \operatorname{rank}(\Sigma_L(\vartheta^*) D_{\varphi'(\pi^*)} \Sigma_L(\vartheta^*))$, where Z_i are independent standard normal variables and $\xi_i(\vartheta^*)$ are the nonzero eigenvalues of

$$D_{\varphi'(\pi^*)}\left(D_{\pi^*} - \pi^*\pi^{*t} - (D_{\pi^*})^{\frac{1}{2}} A_L(A_L^t A_L)^{-1} A_L^t (D_{\pi^*})^{\frac{1}{2}}\right).$$

c) It is straightforward from Corollary 2.1, since the maximum likelihood estimate equals the minimum K_ϕ-divergence estimate when $\phi = \phi_\tau$, with $\tau = 1$. □

3. APPLICATIONS TO TESTING COMPOSITE NULL HYPOTHESES

Using Theorems 2.2 and 2.3 for large n and significance level α, we reject the null hyphotesis given in (2.1) if $nK_{\phi_A}(X/n, \widehat{\pi}) > t_\alpha$ where t_α satisfies the condition

$$\sup_{\vartheta\in\Theta_0} P\left(\sum_{i=1}^{r} \xi_i(\vartheta) Z_i^2 > t_\alpha\right) \leq \alpha$$

for independent standard normal variables Z_i, $r = \text{rank}\,(\Sigma(\vartheta)D_{\varphi'_A(\pi^*)}\Sigma(\vartheta))$, and nonzero eigenvalues $\xi_i(\vartheta)$ of $D_{\varphi'_A(\pi^*)}\Sigma(\vartheta)$, where $\Sigma(\vartheta) = \Sigma_D(\vartheta)$ if $\hat{\pi} = \hat{\pi}_{\phi_B}$, is the minimum K_ϕ-divergence estimate or $\Sigma(\vartheta) = \Sigma_L(\vartheta)$, if $\hat{\pi} = \tilde{\pi}$ is the maximum likelihood estimate.

From a practical point of view we have two ways to carry out the test:

a) Given ϑ fixed we can find the value $t_\alpha(\vartheta)$ verifying

$$P\left(\sum_{i=1}^{r} \xi_i(\vartheta) Z_i^2 > t_\alpha(\vartheta)\right) \leq \alpha$$

and then we can compute $t_\alpha = \sup_{\vartheta\in\Theta_0} t_\alpha(\vartheta)$.

b) Given a value of the statistic we can calculate for each ϑ

$$p(\vartheta) = P\left(\sum_{i=1}^{r} \xi_i(\vartheta) Z_i^2 > nK_{\phi_A}(X/n, \tilde{\pi})\right)$$

and if $\sup_{\vartheta\in\Theta_0} p(\vartheta) < \alpha$ the we have evidence to reject the null hypothesis.

In the above theorems we have obtained the asymptotic distribution of all member of the statistic family K_ϕ, under composite null hypotheses, which has been necessary to built the corresponding test. Now we calculate the asymptotic distribution of $nK_{\phi_A}(X/n, \hat{\pi}_{\phi_B})$ when null hypotheses is not true so that it will allow us to determinate the asymptotic power of this tests.

Theorem 3.1. Let $\pi_1 = f(\vartheta^1)$ be the point in which we want to determinate the power. Assume that there exists $\pi^* = f(\vartheta^*)$ such that $\hat{\pi}_{\phi_B} \to \pi^*$,

$$\left(\vartheta^* = \arg\min_{\vartheta\in\Theta} K_\phi\left(f(\vartheta^1), f(\vartheta)\right),\ f(\vartheta^1) \neq f(\vartheta^*)\right)$$

and that under alternative hypotheses

$$\sqrt{n}\left((X/n, \hat{\pi}_{\phi_B}) - (\pi_1, \pi^*)\right) \xrightarrow[n\to\infty]{L} N(0, \Sigma)$$

with $\Sigma = \begin{pmatrix} \Sigma_{\pi_1} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ for $\Sigma_{\pi_1} = D_{\pi^1} - \pi_1\pi_1^t$, $\Sigma_{21} = \Sigma_{12}$ and for one unknown matrix Σ_{22} which depends on the model under consideration. Then

$$\sqrt{n}\left(K_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) - K_{\phi_A}(\pi_1, \pi^*)\right) \xrightarrow[n\to\infty]{L} N(0, \sigma^2(\pi_1))$$

with $\sigma^2(\pi_1) = T_1^t \Sigma_{\pi^1} T_1 + T_2^t \Sigma_{22} T_2 + 2T_1^t \Sigma_{12} T_2$ where

$$T_1 = (t_1^1, ..., t_m^1)^t, / t_i^1 = \left(\frac{\partial K_\phi(p,q)}{\partial p_i}\right)_{(\pi_1,\pi^*)}$$

$$T_2 = (t_1^2, ..., t_m^2)^t, / t_i^2 = \left(\frac{\partial K_\phi(p,q)}{\partial q_i}\right)_{(\pi_1,\pi^*)}$$

Proof. Expanding in a Taylor's series $K_{\phi_A}(X/n, \hat{\pi}_{\phi_B})$ around

$$(\pi_1, \pi^*) = (\pi_{11}, \pi_{12} ..., \pi_{1m}, \pi_1^*, \pi_2^* ..., \pi_m^*)$$

we have

$$K_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) = K_{\phi_A}(\pi_1, \pi^*) + \sum_{i=1}^{m} \left(t_i^1 \left(X_i/n - \pi_{1i}\right) + t_i^2 \left(\hat{\pi}_{i\phi_B} - \pi_i^*\right)\right) + o_p(n^{-1})$$

where $t_i^1 = \left(\frac{\partial K_\phi(p,q)}{\partial p_i}\right)_{(\pi_1,\pi^*)}$ and $t_i^2 = \left(\frac{\partial K_\phi(p,q)}{\partial q_i}\right)_{(\pi_1,\pi^*)}$.

Denoting $T_1 = (t_1^1, ..., t_m^1)^t$ and $T_2 = (t_1^2, ..., t_m^2)^t$, we get

$$K_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) = K_{\phi_A}(\pi_1, \pi^*) + T_1^t \left(X_i/n - \pi_{1i}\right) + T_2^t \left(\hat{\pi}_{\phi_B} - \pi^*\right) + o_p(n^{-1}).$$

So the asymptotic distribution of the random variables

$$\sqrt{n}\left(K_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) - K_{\phi_A}(\pi_1, \pi^*)\right)$$

and

$$\sqrt{n}\left(T_1^t \left(X_i/n - \pi_{1i}\right) + T_2^t \left(\hat{\pi}_{\phi_B} - \pi^*\right)\right)$$

is the same. As

$$\sqrt{n}\left((X/n, \hat{\pi}_{\phi_B}) - (\pi_1, \pi^*)\right) \xrightarrow[n\to\infty]{L} N(0, \Sigma)$$

with

$$\Sigma = \begin{pmatrix} \Sigma_{\pi_1} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

we have that

$$\sqrt{n}\left(T_1^t \left(X_i/n - \pi_{1i}\right) + T_2^t \left(\hat{\pi}_{\phi_B} - \pi^*\right)\right) \xrightarrow[n\to\infty]{L} N(0, T_1^t \Sigma_{\pi_1} T_1 + T_2^t \Sigma_{22} T_2 + 2T_1^t \Sigma_{12} T_2).$$

□

From Theorem 3.1, the asymptotic power in π_1 is given by

$$\beta_\pi(\pi_1) = P_{\pi_1}(nK_{\phi_A}(X/n, \hat{\pi}_{\phi_B}) > t_\alpha) = 1 - F_{N(0,1)}\left(\frac{t_\alpha - nK_{\phi_A}(\pi_1, \pi^*)}{\sqrt{n}\sigma(\pi_1)}\right)$$

where t_α is the critical value such that

$$\sup_{\vartheta \in \Theta_0} P\left(\sum_{i=1}^{r} \xi_i(\vartheta) Z_i^2 > t_\alpha\right) \leq \alpha.$$

Besides as $\beta_\pi(\pi_1) \xrightarrow[n\to\infty]{} 1$, the contrast is consistent in Fraser sense [7].

Remark 3.1. In order to apply the above test, we have to calculate a probability of a linear combination of chi-squared distributions. These probabilities can be computed using the methods given by Rao and Scott [14]. These authors suggest to consider the following approximate distributions of $\sum_{i=1}^{r} \xi_i(\vartheta^*) Z_i^2$,

i) $\bar{\xi}(\vartheta^*)\chi_r^2$ where $\bar{\xi} = \sum_{i=1}^{r} \xi_i(\vartheta^*)/r$.

ii) $\xi^*(\vartheta^*)\chi_r^2$ where $\xi^*(\vartheta^*) = \max\{\xi_1(\vartheta^*), ..., \xi_r(\vartheta^*)\}$.

iii) $\bar{\xi}(\vartheta^*)(1+\lambda^2)\chi_\nu^2$ where $\nu = \dfrac{r}{1+\lambda^2}$ and $\lambda^2 = \sum_{i=1}^{r} \dfrac{\left(\xi_i(\vartheta^*) - \bar{\xi}(\vartheta^*)\right)^2}{r\bar{\xi}(\vartheta^*)^2}$.

4. NUMERICAL EXAMPLE

We consider a genetic problem in which each individual only can have one of six different genotypes and we want to test the probability of all of them. We collected information from 600 individuals and we classified them according to its genotype so that we build 6 classes.

We test the following hypotheses

$$H_0 : \pi = f(\vartheta) =$$
$$= \left(\vartheta_1^2, \vartheta_2^2, (1-\vartheta_1-\vartheta_2)^2, 2\vartheta_1\vartheta_2, 2\vartheta_1(1-\vartheta_1-\vartheta_2), 2\vartheta_2(1-\vartheta_1-\vartheta_2)\right)^t \quad (4.1)$$

where $\vartheta_1 > 0$, $\vartheta_2 > 0$ and $\vartheta_1 + \vartheta_2 < 1$.

Table 1 shows the observed frequencies in each class

Table 1.

Genotype	1	2	3	4	5	6
Observed Frequencies	30	90	94	98	89	199

Firstly we have to estimate two parameters using the minimum K_ϕ-divergence method. In order to test the hypotheses (4.1) we are going to use the statistic

$$nK_{\phi_2}(X/n, \hat{\pi}_{\phi_2}) = \sum_{i=1}^{6} n\left(X_i/n - \hat{\pi}_i\right)^2 \text{ with } \phi_2 = x^2 - x$$

and $\hat{\vartheta}_{\phi_2} = \left(\hat{\vartheta}_1, \hat{\vartheta}_2\right)^t$ are those values which minimize

$$nK_{\phi_2}(X/n, \pi) = \sum_{i=1}^{6} n\left(X_i/n - f(\vartheta)\right)^2.$$

Then the minimum K_ϕ-divergence estimates are $\hat{\vartheta}_1 = 0.197$, $\hat{\vartheta}_2 = 0.402$, the probability vector $\hat{\pi}_{\phi_2} = (0.038, 0.16, 0.16, 0.158, 0.158, 0.322)^t$ and the statistic takes the value 0.288.

In order to determinate the size α critical region it is necessary to calculate the critical value t_α such that

$$P\left(\sum_{i=1}^{r}\xi_i(\vartheta^*)Z_i^2 > t_\alpha\right) = \alpha.$$

Particularizing in Theorem 2.2 for $\phi = \phi_2$ we obtain that $\xi_i(\vartheta^*)$ are the nonzero eigenvalues of

$$I\left(D_{\pi^*} - \pi^*\pi^{*t} - \left(D_{\pi^*} - \pi^*\pi^{*t}\right)L^t - L\left(D_{\pi^*} - \pi^*\pi^{*t}\right) + L\left(D_{\pi^*} - \pi^*\pi^{*t}\right)L^t\right)I$$

with $L = IA_D(A_D^tA_D)^{-1}A_DI$, A_D is an $m \times s$ matrix such that

$$A_{D_{ij}} = \left(\frac{\partial f_i(\vartheta^*)}{\partial \vartheta_j}\right) \qquad \text{for } i = 1, 2..., 6;\ j = 1, 2.$$

Replacing π^* with $\hat{\pi}_{\phi_2}$ we get

$$\left(D_{\hat{\pi}_{\phi_2}} - \hat{\pi}_{\phi_2}\hat{\pi}_{\phi_2}^t - \left(D_{\hat{\pi}_{\phi_2}} - \hat{\pi}_{\phi_2}\hat{\pi}_{\phi_2}^t\right)L^t - L\left(D_{\hat{\pi}_{\phi_2}} - \hat{\pi}_{\phi_2}\hat{\pi}_{\phi_2}^t\right) + L\left(D_{\hat{\pi}_{\phi_2}} - \hat{\pi}_{\phi_2}\hat{\pi}_{\phi_2}^t\right)L^t\right) =$$

$$= \begin{pmatrix} 0.046 & 0.022 & 0.022 & -0.019 & -0.019 & -0.048 \\ 0.022 & 0.068 & 0.039 & -0.042 & 0.020 & -0.075 \\ 0.022 & 0.039 & 0.069 & 0.019 & -0.054 & -0.077 \\ -0.019 & -0.042 & 0.019 & 0.075 & -0.053 & 0.024 \\ -0.019 & 0.020 & -0.054 & -0.053 & 0.075 & 0.024 \\ -0.048 & -0.075 & -0.077 & 0.024 & 0.024 & 0.117 \end{pmatrix}$$

and the eigenvalues are
$\xi_1(\hat{\vartheta}_{\phi_2}) = 0.0400$; $\xi_2(\hat{\vartheta}_{\phi_2}) = 0.248$; $\xi_3(\hat{\vartheta}_{\phi_2}) = -0.0075$; $\xi_4(\hat{\vartheta}_{\phi_2}) = 0.1635$; $\xi_5(\hat{\vartheta}_{\phi_2}) = 0.0057$; $\xi_6(\hat{\vartheta}_{\phi_2}) = 0.00014$.

From Remark 3.1 it follows that the distribution of $\sum_{i=1}^{r}\xi_i(\theta^*)Z_i^2$ can be approximated by

i) The distribution of $\bar{\xi}(\hat{\vartheta}_{\phi_2})\chi_r^2$ where $\bar{\xi}(\hat{\vartheta}_{\phi_2}) = \sum_{i=1}^{r}\xi_i(\hat{\vartheta}_{\phi_2})/r$.

$$\bar{\xi}(\hat{\vartheta}_{\phi_2}) = 0.075;\ \chi^2_{6,0.05} = 12.6;\ \bar{\xi}(\hat{\vartheta}_{\phi_2})\chi^2_{6,0.05} = 0.945.$$

ii) The distribution of $\xi^*(\hat{\vartheta}_{\phi_2})\chi_r^2$ where $\xi^*(\hat{\vartheta}_{\phi_2}) = \max\left\{\xi_1(\hat{\vartheta}_{\phi_2}), ..., \xi_r(\hat{\vartheta}_{\phi_2})\right\}$.

$$\xi^*(\hat{\vartheta}_{\phi_2}) = 0.248;\ \chi^2_{6,0.05} = 12.6;\ \bar{\xi}(\hat{\vartheta}_{\phi_2})\chi^2_{6,0.05} = 3.12.$$

iii) The distribution of $\bar{\xi}(\hat{\vartheta}_{\phi_2})(1+\lambda^2)\chi_\nu^2$ where $\nu = \dfrac{r}{1+\lambda^2}$ and

$$\lambda^2 = \sum_{i=1}^{r}\frac{\left(\xi_i(\hat{\vartheta}_{\phi_2}) - \bar{\xi}(\hat{\vartheta}_{\phi_2})\right)^2}{r\bar{\xi}(\hat{\vartheta}_{\phi_2})^2},$$

$$\lambda^2 = 1.665;\ \nu = 2.25;\ \chi^2_{2,0.05} = 5.99;\ \bar{\xi}(\hat{\vartheta}_{\phi_2})(1+\lambda^2)\chi^2_\nu = 1.27.$$

Besides we have computed 6 samples with sample sizes equal to 1000, from a χ^2_1 distribution and multiplying each one by the corresponding eigenvalues, we obtained the critical value of size $\alpha = 0.05$ for $\sum_{i=1}^{6} \xi_i(\hat{\vartheta}_{\phi_2})\chi^2_1$, being equal to 1.2711.

If we compare these values with the statistic $nK_{\phi_2}(X/n, \hat{\pi}_{\phi_2}) = 0.288$, in all cases we get the same conclusion, there is no statistic evidence to reject the null hypotheses.

ACKNOWLEDGEMENTS

We thank a referee for a careful examination of the paper and suggestions leading to its improvements. This work was supported by Grant DGI BFM 2003-00892.

(Received November 2, 2000.)

REFERENCES

[1] N. W. Birch: A new proof of the Pearson-Fisher theorem. Ann. Math. Statist. *35*(1964), 817–824.

[2] Y. M. M. Bishop, S. E. Fienberg, and P. W. Holland: Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, MA 1975.

[3] J. Burbea: J-divergences and related topics. Encyclopedia Statist. Sci. *44* (1983), 290–296.

[4] J. Burbea and C. R. Rao: On the convexity of some divergence measures based on entropy functions. IEEE Trans. Inform. Theory *28* (1982), 489–495.

[5] N. Cressie and T. R. C.: Multinomial Goodness-of-fit Tests. J. Roy. Statist. Soc. Ser. B *46* (1984), 440–464.

[6] I. Csiszár: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hungar. *2* (1967), 299–318.

[7] D. A. S. Fraser: Nonparametric methods in Statistics. Wiley, New York 1957.

[8] J. C. Fryer and C. A. Robertson: A comparison of some methods for estimating mixed normal distributions. Biometrika *59* (1972), 639–648.

[9] D. Morales, L. Pardo, and I. Vajda: Asymptotic divergence of estimates of discrete distributions. J. Statist. Plann. Inference *48* (1995), 347–369.

[10] M. C. Pardo: On Burbea-Rao divergence based goodness-of-fit tests for multinomial models. J. Multivariate Anal. *69* (1999), 65–87.

[11] T. Pérez and J. A. Pardo: The K_ϕ-divergence statistics for categorical data problem. Submitted .

[12] T. Pérez and J. A. Pardo: Minimum K_ϕ-divergence estimator. Appl. Math. Lett. (to appear).

[13] C. R. Rao: Diversity and dissimilarity coefficients: a unified approach. J. Theoret. Population Biology *21* (1982), 24–43.

[14] J. N. K. Rao and A. J. Scott: The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables. J. Amer. Statist. Assoc. *76* (1981), 221–230.

[15] T. R. C. Read and N. A. C. Cressie: Goodness of fit Statistics for discrete multivariate data. Springer–Verlag, Berlin 1988.

[16] T. Vajda and K. Vašek: Majorization, concave entropies, and comparison of experiments. Prob. Control Inform. Theory *14* (1985), 105–115.

[17] K. Zografos, K. Ferentinos, and T. Papaioannou: φ-divergence statistics: sampling properties and multinomial goodness of fit divergence tests. Comm. Statist. Theory Methods *19* (1990), 5, 1785–1802.

Prof. Dr. Julio A. Pardo, Department of Statistics and O.R. (I), Faculty of Mathematics, Complutense University of Madrid, 28040 Madrid. Spain.
e-mail: julio_pardo@mat.ucm.es

Dr. Teresa Pérez, Department of Statistics and O.R. (III), School of Statistics, Complutense University of Madrid, 28040 Madrid. Spain.
e-mail: teperez@estad.ucm.es