KYBERNETIKA — VOLUME 36 (2000), NUMBER 4, PAGES 437-454

SECOND ORDER ASYMPTOTIC DISTRIBUTION OF THE R_ϕ–DIVERGENCE GOODNESS–OF–FIT STATISTICS[1]

María Carmen Pardo

The distribution of each member of the family of statistics based on the R_ϕ-divergence for testing goodness-of-fit is a chi-squared to $o(1)$ (Pardo [12]). In this paper a closer approximation to the exact distribution is obtained by extracting the ϕ-dependent second order component from the $o(1)$ term.

1. INTRODUCTION

For a sequence of n observations on a multinomial random vector $X = (X_1, \dots, X_M)^t$ with probability vector $\pi = (\pi_1, \dots, \pi_M)^t$, $\sum_{i=1}^M \pi_i = 1$. Let $\pi_0 = (\pi_{01}, \dots, \pi_{0M})^t$ a prespecified probability vector with $\pi_{0i} > 0$ for each i and $\sum_{i=1}^M \pi_{0i} = 1$. Then to test the simple hypothesis $H_0 : \pi = \pi_0$ against $H_1 : \pi \neq \pi_0$, the most commonly used statistic is Pearson's X^2 (Pearson [14]);

$$X^2 = \sum_{i=1}^{M} \frac{(X_i - n\pi_{0i})^2}{n\pi_{0i}}$$

which is asymptotically distributed as a chi-squared with $M-1$ degrees of freedom.

Cressie and Read [7] and Read and Cressie [20] introduced the power divergence family of statistics

$$2nI^\lambda(X/n, \pi_0) = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{M} X_i \left(\left(\frac{X_i}{n\pi_{0i}} \right)^\lambda - 1 \right), \quad -\infty < \lambda < \infty$$

where the index parameter $\lambda \in R$, $\lambda \neq -1, 0$. It can be easily seen that Pearson's X^2 $(\lambda = 1)$, the loglikelihood ratio statistic $(\lambda \to 0)$, the Freeman–Tukey statistic $(\lambda = -1/2)$, the modified loglikelihood ratio statistic $(\lambda \to -1)$ and the Neyman modified X^2 $(\lambda = -2)$, are all special cases of this family. These authors proved that under the same regularity conditions each member of the power divergence family follows the same asymptotic distribution (a chi-squared with $M-1$ degrees of freedom).

[1]This work was supported by Grant DGICYT PB96-0635.

The classical chi-square approximation of the distribution of Pearson's statistic X^2 is suspected to be poor unless each expected cell frequency is reasonably large. A common rule of thumb used by statisticians for many years is that for the chi-square approximation to be meaningful, each expected cell frequency must be more than 5 (Rao [16], p. 396). To circumvent this difficulty various approximations to the distribution of the statistics $2nI^{\lambda}(X/n, \pi_0)$ have been given by Read and Cressie [20] and a comprehensive review of the most important results in this area can be found in their book as well as the references therein. One of these approximations is given using local Edgeworth expansions. This approximation was firstly given by Yarnold [25] for the Pearson chi-square statistic under the null hypothesis, by Siotani and Fujikoshi [22] for the log-likelihood ratio and Freeman–Tukey statistics and for the power divergence family of statistics by Read and Cressie [20]. We therefore omit motivation or justification of this method of asymptotic expansion for the distribution function of a quadratic form.

We consider the family of statistics based on the R_ϕ-divergence measure between the observed proportions x/n and the hypothesized proportions π_0 introduced by Pardo [12]

$$S_\phi(X/n, \pi_0) = -\frac{M}{\phi''(1/M)} 8nR_\phi(X/n, \pi_0),$$

where

$$R_\phi(X/n, \pi_0) = \sum_{i=1}^{M} \left\{ \phi\left(\frac{X_i/n + \pi_{0i}}{2}\right) - \frac{1}{2}\left[\phi\left(X_i/n\right) + \phi\left(\pi_{0i}\right)\right] \right\}$$

for a given continuous concave function $\phi : (0, \infty) \to R$ with $\phi(0) = \lim_{t\downarrow 0} \phi(t) \in (-\infty, \infty]$. The R_ϕ-divergence was introduced and studied by Rao [17], Burbea and Rao ([4, 5]), Burbea [3] in many statistical problems. Some properties of this family of divergences can be seen in Pardo and Vajda [13]. If the R_ϕ-divergence is "too large" the null hypothesis is rejected. An approximation to the exact distribution of the statistic $S_\phi(X/n, \pi_0)$ under uniform hypothesis was obtained from

$$T_E(c) = P\left(S_\phi(X/n, \pi_0) < c\right) = T_{\chi_{M-1}}(c) + o(1) \tag{1}$$

where T_{χ_ν} is the chi-square distribution function on ν degrees of freedom. This result holds for every member of the family, as $n \to \infty$.

Several reasons justify the choice of uniform hypothesis. Sturges [23] initiated the study of the choice of cells and recommended that the cell would be chosen to have equal probabilities with $M = 1 + 2.303 \log_{10} n$. Mann and Wald [11] for a sample size n (large) and a significance level γ, recommended $M = 4\left(\frac{2n^2}{z_\gamma^2}\right)^{1/5}$ where z_γ is the upper γ—point of the standard normal distribution. Schorr [21], confirmed that the 'optimum' M is smaller than the value given by Mann and Wald and suggested to use $M = 2n^{2/5}$. In Greenwood and Nikolin [9] it is suggested to use $M \leq \min\left(\frac{1}{\gamma}, \log n\right)$. Cohen and Sackrowitz [6] proved that for the above hypothesis a critical region of the form $\sum_{i=1}^{M} h_i(x_i) > c$, where c is a positive constant, h_i, $i = 1, \ldots, M$, are convex functions and $x_i \geq 0$, $i = 1, \ldots, M$, is unbiased. In our case if we choose ϕ so that S_ϕ was convex, the proposed tests are unbiased for equal cell probabilities. Bednarski and Ledwina [2] established that for every fixed number of observations, every continuous and reflexive function $h : \Delta_M \times \Delta_M \to R^+$ where $\Delta_M = \{(p_1, \ldots, p_M)^t / \sum_{i=1}^M p_i = 1, p_i \geq 0, \ i = 1, \ldots, M\}$ and every $0 < c < \sup\{c/P(h(p,x) \geq c) < 1, p \in \Delta_M\}$, exists $q \in \Delta_M$ such that the test with critical region $h(q,x) > c$ is biased for testing $H_0 : p = q$. In the book of Read and Cressie [20] (pp. 148–150) an important historical perspective illustrating the importance of choosing equiprobable cells can be seen. All this justifies that in the rest of the paper we consider equiprobable cells. The statistic, $S_\phi(X/n, \pi_0)$ that we study in this paper is a continuous function in $\Delta_M \times \Delta_M - \{(0,0)\}$ and then, it is not unbiased in general for unequal cell probabilities case.

In this paper following Read [19] we extract the ϕ dependent second order component from the $o(1)$ term in (1) to obtain an asymptotic expansion for the distribution function of the quadratic form given by the statistic $S_\phi(X/n, \pi_0)$ closer to the exact distribution function than (1).

The power divergence family of statistics is a particular case of the family of goodness-of-fit statistics studied by Zografos et al [26] based on the measure of divergence called φ-divergence, introduced by Csiszár [8] and Ali and Silvey [1] is given by

$$C_\varphi(X/n, \pi_0) = \frac{2n}{\varphi''(1)} \sum_{i=1}^M \frac{X_i}{n} \varphi\left(\frac{X_i}{n\pi_{0i}}\right)$$

for any continuous convex function $\varphi : [0, \infty) \to R \cup \{\infty\}$, where $0\varphi(0/0) = 0$ and $0\varphi(p/0) = \lim_{u\to\infty} \frac{\varphi(u)}{u}$ (for a systematic theory of these divergences see Liese and Vajda [10] and Vajda [24]). We can observe that for $\pi_{0i} = 1/M$, $i = 1, \ldots, M$,

$$S_\phi(X/n, \pi_0) = C_{\varphi_\phi}(X/n, \pi_0)$$

for

$$\varphi_\phi(t) = \phi\left(\frac{\frac{t}{M} + \frac{1}{M}}{2}\right) - \frac{1}{2}\phi\left(\frac{t}{M}\right) - \frac{1}{2}\phi\left(\frac{1}{M}\right)$$

when φ_ϕ is convex.

So the results obtained in this paper are valid for some of the families of statistics C_{φ_ϕ}.

For example an important family of statistics S_ϕ is obtained if we consider the family of functions,

$$\phi(x) = \phi_\alpha(x) = \begin{cases} (1-\alpha)^{-1}(x^\alpha - x) & \alpha \neq 1 \\ -x \log x & \alpha = 1. \end{cases}$$

In this case the S_{ϕ_α} is convex if and only if $\alpha \in [1,2]$, for $M > 2$, and if only if $\alpha \in [1,2]$ or $\alpha \in [3, 11/3]$, for $M = 2$. Moreover,

$$C_{\varphi_\alpha}(X/n, \pi_0) = S_{\phi_\alpha}(X/n, \pi_0)$$

for

$$\varphi_\alpha(x) = \begin{cases} \frac{-2(\frac{1}{2}x+\frac{1}{2})^\alpha + x^\alpha + 1}{2M^\alpha \alpha(\alpha-1)} & \alpha > 0, \alpha \neq 1 \\ \frac{1}{2M}\left(x \log \frac{2x}{x+1} + \log \frac{2}{x+1}\right) & \alpha = 1. \end{cases}$$

Note that

$$S_{\phi_2}(X/n, \pi_0) = \sum_{i=1}^{M} \frac{(X_i - n/M)^2}{n/M}.$$

Then the result obtained by Yarnold [25] appears as special case of our main theorem.

2. NOTATION AND PRELIMINARY RESULTS

Define $W_j = \sqrt{n}(X_j/n - \pi_{0j})$, with $\pi_{0j} = 1/M$, $j = 1, \ldots, M$ and let $W = (W_1, \ldots, W_r)^t$ where $r = M - 1$. Therefore, W is a lattice random vector taking values in the lattice

$$L = \left\{w = (w_1, \ldots, w_r)^t : w = \sqrt{n}(x/n - \pi_0^*) \text{ and } x \in K\right\}, \tag{2}$$

where

$$\pi_0^* = (\pi_{01}, \ldots, \pi_{0r})^t$$

and

$$K = \left\{x = (x_1, \ldots, x_r)^t : x_j \geq 0 \text{ integer}, \ j = 1, \ldots, r; \ \sum_{j=1}^{r} x_j \leq n\right\}.$$

The asymptotic expansion of the random vector W (Siotani and Fujikoshi [22]) is given by

$$P(W = w) = n^{-r/2}\varphi(w)\left\{1 + n^{-1/2}h_1(w) + n^{-1}h_2(w) + O(n^{-3/2})\right\} \tag{3}$$

where

$$\varphi(w) = (2\pi)^{-r/2}|\Omega|^{-1/2}\exp\left(-\frac{1}{2}w^t\Omega^{-1}w\right)$$

is the multivariate Normal density function, and

$$\begin{aligned} h_1(w) &= -\frac{1}{2}\sum_{j=1}^{M}\frac{w_j}{\pi_{0j}} + \frac{1}{6}\sum_{j=1}^{M}\frac{w_j^3}{\pi_{0j}^2}, \\ h_2(w) &= \frac{1}{2}(h_1(w))^2 + \frac{1}{12}\left(1 - \sum_{j=1}^{M}\frac{1}{\pi_{0j}}\right) + \frac{1}{4}\sum_{j=1}^{M}\frac{w_j^2}{\pi_{0j}^2} - \frac{1}{12}\sum_{j=1}^{M}\frac{w_j^4}{\pi_{0j}^3} \end{aligned} \tag{4}$$

with

$$w_M = -\sum_{j=1}^{r} w_j, \ \Omega = \operatorname{diag}(\pi_0^*) - \pi_0^*\pi_0^{*t}.$$

This result gives us a local Edgeworth approximation for the probability of W at each point $w \in L$. In the case where W has a continuous probability distribution function, we have that

$$P(W \in B) = \int_B \cdots \int \varphi(w)\{1 + n^{-1/2}h_1(w) + n^{-1}h_2(w)\}\, \mathrm{d}w + O(n^{-3/2}).$$

However, when W has a lattice distribution, as occurs here, then Yarnold [25] indicated that the above expansion is not valid. Rao [15] expressed this lattice sum as a Stieltjes integral when B is a Borel set. However Rao's expansion is difficult to apply and Yarnold has obtained a useful evaluation for the case when B is an extended convex set, i.e., when B is a set which can be represented as

$$\begin{aligned} B = \{ & w = (w_1, \ldots, w_r)^t : \gamma_s(w^*) < w_s < \theta_s(w^*), \\ & w^* = (w_1, \ldots, w_{s-1}, w_{s+1}, \ldots, w_r)^t \in B_s \} \end{aligned} \tag{5}$$

where $B_s \subset R^{r-1}$ and γ_s, θ_s are continuous functions on R^{r-1}, $s = 1, \ldots, r$, which is given by

$$P(W \in B) = J_1 + J_2 + J_3 + O(n^{-3/2})$$

where

$$\begin{aligned} J_1 = \;& \int_B \cdots \int \varphi(w)\{1 + n^{-1/2}h_1(w) + n^{-1}h_2(w)\}\, \mathrm{d}w, \\ J_2 = \;& -n^{-1/2} \sum_{s=1}^{r} n^{-(r-s)/2} \sum_{w_{s+1} \in L_{s+1}} \cdots \sum_{w_r \in L_r} \int_{B_s} \cdots \int \\ & \left(S_1\left(\sqrt{n} w_s + n\pi_{0s}\right) \varphi(w)\right)_{\gamma_s(w^*)}^{\theta_s(w^*)} \mathrm{d}w_1 \ldots \mathrm{d}w_{s-1}, \\ J_3 = \;& O(n^{-1}), \end{aligned}$$

with h_1 and h_2 defined in (4),

$$\begin{aligned} L_j &= \{w_j : w_j = \sqrt{n}(x_j/n - \pi_{0j}) \text{ and } x_j \text{ is integer}\}, \\ S_1(t) &= t - [t] - 1/2, \end{aligned}$$

$\theta_s(w^*)$ and $\gamma_s(w^*)$ are as in (5), and

$$\begin{aligned} h(w)_{\gamma_s(w^*)}^{\theta_s(w^*)} = \;& h(w_1, \ldots, w_{s-1}, \theta_s(w^*), w_{s+1}, \ldots, w_r) \\ & - h(w_1, \ldots, w_{s-1}, \gamma_s(w^*), w_{s+1}, \ldots, w_r). \end{aligned}$$

3. THE EXPANSION FOR S_ϕ

The general distribution function of the family $S_\phi(X/n, \pi_0)$ under the uniform hypothesis, can be described as follows

$$P(S_\phi(X/n, \pi_0) < c) = P(W \in B_\phi(c))$$

where

$$B_\phi(c) = \{w = (w_1, \ldots, w_r)^t : S_\phi((x/n, x_M/n), \pi_0) < c\}$$

being

$$w_M = -\sum_{j=1}^{r} w_j, \; x = \sqrt{n}w + n\pi_0^* \quad \text{and} \quad x_M = \sqrt{n}w_M + n/M.$$

$B_\phi(c)$ can readily be seen to be an extended convex set where $\gamma_s(w^*)$ and $\theta_s(w^*)$ are chosen such that if $w_s = \gamma_s(w^*)$ or $w_s = \theta_s(w^*)$, $s = 1, \ldots, r$, then it holds $S_\phi((x/n, x_M/n), \pi_0) = c$. Therefore using the result of Yarnold [25] with $B = B_\phi(c)$, the second order expansion for the distribution function of the general family S_ϕ is obtained in the following theorem.

Theorem 1. Let $\phi : (0, \infty) \to R$ a concave and twice continuously differentiable function with $\phi''(1/M)$ negative. The asymptotic expansion for the distribution function of the statistic $S_\phi(X/n, \pi_0)$ can be expressed as

$$P(S_\phi(X/n, \pi_0) < c) = J_1^\phi + J_2^\phi + J_3^\phi + O(n^{-3/2})$$

where J_1^ϕ, J_2^ϕ and J_3^ϕ are defined by J_1, J_2 and J_3 respectively from Yarnold's result [25] by setting $B = B_\phi(c)$. Furthermore

$$J_1^\phi = P(\chi_r^2 < c) + \frac{(M-1)}{96n}\Bigg\{P(\chi_r^2 < c)(-8(M+1)) + P(\chi_{r+2}^2 < c)\Bigg(-\frac{21}{M^2} \times \frac{\phi^{IV}(1/M)}{\phi''(1/M)}(M-1) + \frac{18}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 (M-2) + 24M\Bigg) + P(\chi_{r+4}^2 < c) \times \Bigg(-\frac{24\phi'''(1/M)}{M\phi''(1/M)}(M-2) + \frac{21\phi^{IV}(1/M)}{M^2\phi''(1/M)}(M-1) - \frac{36}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 (M-2) - 24(M-1)\Bigg) + P(\chi_{r+6}^2 < c)\Bigg(2(M-2)\Bigg(\frac{12\phi'''(1/M)}{M\phi''(1/M)} + \frac{9}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 + 4\Bigg)\Bigg)\Bigg\} + O(n^{-3/2})$$

Also J_2^ϕ can be approximated to first order by

$$\hat{J}_2^\phi = (N_\phi(c) - n^{r/2}V_\phi(c))e^{-c/2} \Big/ \left((2\pi n)^{r/2} M^{-M/2}\right),$$

where

$$N_\phi(c) = \text{the number of lattice points in } B_\phi(c)$$

and

$$
\begin{aligned}
V_\phi(c) = & \text{ the volume of } B_\phi(c) \\
= & \frac{(\pi c)^{r/2}}{\Gamma(1+r/2)}\left(\frac{1}{M}\right)^{M/2}\left\{1+\frac{c(M-1)}{32M^2(M+1)n}\left(\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2(6(M-2))\right.\right. \\
& \left.\left. -\frac{7\phi^{IV}(1/M)}{\phi''(1/M)}(M-1)\right)\right\} + O(n^{-3/2}).
\end{aligned}
$$

Proof. The proof is completed in two parts and the results are derived in a similar fashion to that for the power divergence statistic by Read [19].

Firstly J_1^ϕ is evaluated for which we consider the transformation

$$
z^t = w^t H = w^t(I_r, -\mathbf{1})\,\mathrm{diag}(\pi_0)^{-1/2}A \tag{6}
$$

where

I_r is the identity matrix of order $r = M-1$,

$\mathbf{1} = (1,\ldots,1)^t$ is a $1 \times r$ vector,

$A^t = (a_1,\ldots,a_M)$ is an $r \times M$ matrix such that $(A,\sqrt{\pi_0})$ is orthogonal

and

$$
\sqrt{\pi_0} = \left(\sqrt{1/M},\ldots,\sqrt{1/M}\right)^t.
$$

On one hand, on being $(A,\sqrt{\pi_0})$ an orthogonal matrix we have that $A^tA = I_r$ and $A^t\sqrt{\pi_0} = 0$. Therefore, as $z^t = w^t(I_r,-\mathbf{1})\,\mathrm{diag}(\pi_0)^{-1/2}A$, it follows that

$$
Az = \mathrm{diag}(\pi_0)^{-1/2}(I_r,-\mathbf{1})^t w = \left(w_1\sqrt{M},\ldots,w_M\sqrt{M}\right)^t.
$$

Consequently $w_j = \sqrt{1/M}a_j^t z$. On the other hand,

$$
H^t\Omega H = A^tA - A^t\sqrt{\pi_0}\sqrt{\pi_0^t}A
$$

and applying that $(A,\sqrt{\pi_0})$ is orthogonal we have that $H^t\Omega H = I_r$. So (3) can be expressed as

$$
P(W=w) = n^{-r/2}|\Omega|^{-1/2}\left\{f(z) + O(n^{-3/2})\right\}
$$

where

$$
f(z) = (2\pi)^{-r/2}\exp\left(-\frac{1}{2}z^tz\right)\left(1+n^{-1/2}g_1(z)+n^{-1}g_2(z)\right) \tag{7}
$$

with

$$
\begin{aligned}
g_1(z) &= -T_1/2 + T_3/6, \\
g_2(z) &= g_1^2(z)/2 + (1-M^2)/12 + T_2/4 - T_4/12
\end{aligned}
$$

and

$$T_1 = \sum_{j=1}^{M}(a_j^t z)\sqrt{M}, \quad T_2 = M\sum_{j=1}^{M}(a_j^t z)^2,$$

$$T_3 = \sum_{j=1}^{M}(a_j^t z)^3\sqrt{M}, \quad T_4 = M\sum_{j=1}^{M}(a_j^t z)^4.$$

From Yarnold's result [25] and (7) it follows that J_1^{ϕ} can be rewritten as

$$J_1^{\phi} = \int_{B_{\phi}^*(c)} \cdots \int f(z)\,\mathrm{d}z$$

where

$$B_{\phi}^*(c) = \left\{z : z^t = w^t H \text{ and } w \in B_{\phi}(c)\right\}. \tag{8}$$

By interpreting $f(z)$ as the continuous density function of a random variable Z, it is possible to interpret J_1^{ϕ} as the distribution function of $S_{\phi}((z^t H^{-1}/\sqrt{n})^t + \pi_0, \pi_0)$ which will be abbreviated $S_{\phi}(z^t H^{-1})$ and its characteristic function is given by

$$c(t) = \int_{R^r} \cdots \int \exp\left(it S_{\phi}(z^t H^{-1})\right) f(z)\,\mathrm{d}z.$$

The function $S_{\phi}(z^t H^{-1})$ can be expanded in a Taylor series as

$$S_{\phi}(z^t H^{-1}) = z^t z + n^{-1/2}\frac{\phi'''(1/M)}{2M\phi''(1/M)}T_3 + n^{-1}\frac{7\phi^{IV}(1/M)}{48M^2\phi''(1/M)}T_4 + O(n^{-3/2}). \tag{9}$$

Furthermore, on being

$$\exp\left(\alpha + n^{-1/2}\beta + n^{-1}\gamma\right) = e^{\alpha}\left(1 + n^{-1/2}\beta + n^{-1}(\gamma + \beta^2/2)) + O(n^{-3/2}\right),$$

it follows that

$$\exp\left(it S_{\phi}(z^t H^{-1})\right) f(z) = (2\pi)^{-r/2}\exp\left(itz^t z - \frac{1}{2}z^t z + n^{-1/2}\frac{\phi'''(1/M)}{2M\phi''(1/M)}T_3 it\right.$$
$$\left. + n^{-1}\frac{7\phi^{IV}(1/M)}{48M^2\phi''(1/M)}T_4 it + O(n^{-3/2})\, it\right)(1 + n^{-1/2}g_1(z)$$
$$+ n^{-1}g_2(z)) = (2\pi)^{-r/2}\exp((2it-1)z^t z/2)\left(1 + n^{-1/2}v_1(z)\right.$$
$$\left. + n^{-1}v_2(z)\right)(1 + n^{-1/2}g_1(z) + n^{-1}g_2(z)) + O(n^{-3/2})$$

where

$$v_1(z) = \frac{\phi'''(1/M)}{2M\phi''(1/M)}T_3 it$$

and

$$v_2(z) = \frac{7\phi^{IV}(1/M)}{48M^2\phi''(1/M)}T_4 it - \frac{1}{8M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 T_3^2 t^2.$$

So it follows that

$$c(t) = \sigma^r E[b(Z)] + O(n^{-3/2})$$

where

$$Z \simeq N(0, \sigma^2 I_r),$$
$$\sigma^2 = (-2it + 1)^{-1}$$

and

$$b(z) = 1 + n^{-1/2}\left(-T_1/2 + T_3/6 + it\frac{\phi'''(1/M)}{2M\phi''(1/M)}T_3\right)$$
$$+ n^{-1}\left(it\frac{\phi'''(1/M)}{12M\phi''(1/M)}(T_3^2 - 3T_1T_3) + \frac{7\phi^{IV}(1/M)}{48M^2\phi''(1/M)}T_4 it\right.$$
$$\left.- \frac{1}{8M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 T_3^2 t^2 + (-T_1/2+T_3/6)^2/2 + (1-M^2)/12 + T_2/4 - T_4/12\right).$$

On being

$$AZ \simeq N(0, \sigma^2 AA^t)$$

with

$$AA^t = \begin{pmatrix} 1-\frac{1}{M} & \dots & -\frac{1}{M} \\ \dots & \dots & \dots \\ -\frac{1}{M} & \dots & 1-\frac{1}{M} \end{pmatrix},$$

$(a_k Z^t, a_j Z^t)$ is a bidimensional normal with mean vector $(0,0)^t$ and variance-covariance matrix $\sigma^2\begin{pmatrix} 1-1/M & -1/M \\ -1/M & 1-1/M \end{pmatrix}$. So the random variable $a_k Z^t$ conditioned by $a_j Z^t = t$ is a Normal with mean $-t/(M-1)$ and variance $(1-1/M)(1-(1/(M-1)^2)$. Noting that if X is a Normal with mean μ and standard deviation σ, then

$$E[(X-\mu)^r] = \begin{cases} 0 & r \text{ odd} \\ \frac{r!\sigma^r}{(r/2)!2^{r/2}} & r \text{ even} \end{cases}.$$

it follows that

$$E[(a_j^t Z)] = E[(a_j^t Z)^3] = 0,$$

$$E[(a_k^t Z)(a_j^t Z)] = \begin{cases} -\frac{\sigma^2}{M} & k \neq j \\ \sigma^2\left(1-\frac{1}{M}\right) & k = j \end{cases}$$

since for $k \neq j$

$$\begin{aligned} E[(a_k^t Z)(a_j^t Z)] &= E\left[E[(a_k^t Z)(a_j^t Z)|a_k^t Z = t]\right] = E\left[(a_k^t Z)E[a_j^t Z|a_k^t Z = t]\right] \\ &= E[(a_k^t Z)(-1)(a_k^t Z)/(M-1)] = -E[(a_k^t Z)^2]/(M-1) = -\frac{\sigma^2}{M} \end{aligned}$$

and for $k = j$, it is clear that the expectation is given by

$$E[(a_j^t Z)^2] = \sigma^2(1 - \frac{1}{M}).$$

Analogously,

$$E[(a_k^t Z)^3(a_j^t Z)] = \begin{cases} -\frac{3\sigma^4}{M}\left(1 - \frac{1}{M}\right) & k \neq j \\ 3\sigma^4\left(1 - \frac{1}{M}\right)^2 & k = j \end{cases}$$

$$E[(a_k^t Z)^2(a_j^t Z)^2] = \begin{cases} \sigma^4\left(1 - \frac{2}{M} + \frac{3}{M^2}\right) & k \neq j \\ 3\sigma^4\left(1 - \frac{1}{M}\right)^2 & k = j \end{cases}$$

$$E[(a_k^t Z)^3(a_j^t Z)^3] = \begin{cases} -\sigma^6\left(9\left(1 - \frac{1}{M}\right)^2\frac{1}{M} + \frac{6}{M^3}\right) & k \neq j \\ 15\sigma^6\left(1 - \frac{1}{M}\right)^3 & k = j \end{cases}$$

Then,

$$\begin{aligned} E[T_1] &= E[T_3] = 0, \\ E[T_1^2] &= M\sum_{k=1}^{M}\sum_{j=1}^{M} E[(a_k^t Z)(a_j^t Z)] = M\left(-\frac{\sigma^2}{M}(M-1)+\sigma^2 M\left(1-\frac{1}{M}\right)\right) = 0, \\ E[T_2] &= \sigma^2 M(M-1), \\ E[T_1T_3] &= 0, \\ E[T_4] &= 3\sigma^4(M-1)^2, \\ E[T_3^2] &= 3\sigma^6(2M^2 - 6M + 4) \end{aligned}$$

and hence,

$$\begin{aligned} c(t) &= \sigma^r E[b(Z)] + O(n^{-3/2}) \\ &= \sigma^r + \frac{\sigma^2}{n}\left\{\left(\frac{1-\sigma^{-2}}{2}\right)\frac{\phi'''(1/M)}{12M\phi''(1/M)}\left(3\sigma^6(2M^2 - 6M + 4)\right)\right. \\ &\quad + \frac{7\phi^{IV}(1/M)}{48M^2\phi''(1/M)}\left(\frac{1-\sigma^{-2}}{2}\right)3\sigma^4(M-1)^2 \\ &\quad + \frac{1}{8M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2\left(\frac{1+\sigma^{-4}-2\sigma^{-2}}{4}\right)3\sigma^6(2M^2 - 6M + 4) \\ &\quad \left.+ \frac{1}{12}\sigma^6(M^2-3M+2) + \frac{(1-M^2)}{12} + \frac{\sigma^2 M(M-1)}{4} - \frac{\sigma^4(M-1)^2}{4}\right\} + O(n^{-3/2}) \\ &= \sigma^r + \frac{\sigma^r}{96n}\left\{8(1-M^2) + \sigma^2\left(-\frac{21\phi^{IV}(1/M)}{M^2\phi''(1/M)}(M-1)^2 + (2M^2-6M+4)\frac{9}{M^2}\right.\right. \\ &\quad \left.\times\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 + 24M(M-1)\right) + \sigma^4\left(-\frac{12\phi'''(1/M)}{M\phi''(1/M)}(2M^2 - 6M + 4)\right. \end{aligned}$$

$$+\frac{21\phi^{IV}(1/M)}{M^2\phi''(1/M)}(M-1)^2-\frac{36}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2(M^2-3M+2)-24(M-1)^2\Bigg)$$

$$+\sigma^6\left((2M^2-6M+4)\left(\frac{12\phi'''(1/M)}{M\phi''(1/M)}+\frac{9}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2+4\right)\right)\Bigg\}+O(n^{-3/2}).$$

Taking into account that σ^r is the characteristic function of a χ^2_r and recalling that $c(t)$ is the characteristic function of the distribution J_1^ϕ, we have that

$$J_1^\phi = P(\chi^2_r<c)+\frac{1}{96n}\Bigg\{P(\chi^2_r<c)8(1-M^2)+P(\chi^2_{r+2}<c)\left(-\frac{21\phi^{IV}(1/M)}{M^2\phi''(1/M)}(M-1)^2\right.$$

$$\left.+\frac{9}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2(2M^2-6M+4)+24M(M-1)\right)$$

$$+P(\chi^2_{r+4}<c)\left(-\frac{24\phi'''(1/M)}{M\phi''(1/M)}(M^2-3M+2)+\frac{21\phi^{IV}(1/M)}{M^2\phi''(1/M)}(M-1)^2\right.$$

$$\left.-\frac{36}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2(M^2-3M+2)-24(M-1)^2\right)+P(\chi^2_{r+6}<c)$$

$$\times\left((2M^2-6M+4)\left(\frac{12\phi'''(1/M)}{M\phi''(1/M)}+\frac{9}{M^2}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2+4\right)\right)\Bigg\}+O(n^{-3/2}).$$

Secondly, the approximations for J_2^ϕ and $V_\phi(c)$ are evaluated. On one hand we know that if $w_s=\gamma_s(w^*)$ or $w_s=\theta_s(w^*)$, $s=1,\ldots,r$, then $S_\phi((x/n,x_M/n),\pi_0)=c$ and on the other hand

$$S_\phi\left((x/n,x_M/n),\pi_0\right)=w^t\Omega^{-1}w+o(1),$$

then

$$\varphi(w)=(2\pi)^{-r/2}\,|\Omega|^{-1/2}\exp(-c/2)+o(1)$$

when $w_s=\gamma_s(w^*)$ or $w_s=\theta_s(w^*), s=1,\ldots,r$.

From these facts it follows that

$$\left(S_1(w_s+n\pi_{0s})\,\varphi(w)\right)_{\gamma_s(w^*)}^{\theta_s(w^*)}$$

can be written as

$$(2\pi)^{-r/2}|\Omega|^{-1/2}\exp(-c/2)\left(S_1(\sqrt{n}w_s+n\pi_{0s})\right)_{\gamma_s(w^*)}^{\theta_s(w^*)}+o(1).$$

Therefore by applying Theorem 4 of Yarnold [25] it follows that

$$J_2^\phi=(N_\phi(c)-n^{r/2}V_\phi(c))e^{-c/2}\Big/\left((2\pi n)^r\,|\Omega|\right)^{1/2}+o(1)$$

being $N_\phi(c)$ the number of points in the lattice L which are also in $B_\phi(c)$ and $V_\phi(c)$ the volume of $B_\phi(c)$. So

$$V_\phi(c) = \int_{B_\phi(c)} \cdots \int \mathrm{d}w = |\Omega|^{1/2} \int_{B^*_\phi(c)} \cdots \int \mathrm{d}z$$

where z is defined by (6) and $B^*_\phi(c)$ is defined by (8).

Consider the transformation $z \to u$ such that

$$u^t u = S_\phi(z^t H^{-1}),$$

i. e., from (9)

$$u^t u = z^t z + n^{-1/2}\frac{\phi'''(1/M)}{2M\phi''(1/M)}T_3 + n^{-1}\frac{7\phi^{IV}(1/M)}{48M^2\phi''(1/M)}T_4 + O(n^{-3/2}). \tag{10}$$

By writing

$$z = d_1(u) + n^{-1/2}d_2(u) + n^{-1}d_3(u) + O(n^{-3/2}),$$

(10) can be written as

$$\begin{aligned} u^t u \;=\; & d_1^t(u)d_1(u) + n^{-1/2}\left(2d_1^t(u)d_2(u) + \frac{\phi'''(1/M)}{2M^{1/2}\phi''(1/M)}\sum_{j=1}^{M}(a_j^t d_1(u))^3\right) \\ & + n^{-1}\left(2d_1^t(u)d_3(u) + d_2^t(u)d_2(u) + \frac{3\phi'''(1/M)}{2M^{1/2}\phi''(1/M)}\sum_{j=1}^{M}(a_j^t d_1(u))^2(a_j^t d_2(u))\right. \\ & \left. + \frac{7\phi^{IV}(1/M)}{48M\phi''(1/M)}\sum_{j=1}^{M}(a_j^t d_1(u))^4\right) + O(n^{-3/2}) \end{aligned}$$

where $d_1(u), d_2(u)$ and $d_3(u)$ are such that verify $d_1^t(u)d_1(u) = u^t u$

$$2d_1^t(u)d_2(u) + \frac{\phi'''(1/M)}{2M^{1/2}\phi''(1/M)}\sum_{j=1}^{M}(a_j^t d_1(u))^3 = 0$$

$$\begin{aligned} & 2d_1^t(u)d_3(u) + d_2^t(u)d_2(u) + \frac{3\phi'''(1/M)}{2M^{1/2}\phi''(1/M)}\sum_{j=1}^{M}(a_j^t d_1(u))^2(a_j^t d_2(u)) \\ & \qquad + \frac{7\phi^{IV}(1/M)}{48M\phi''(1/M)}\sum_{j=1}^{M}(a_j^t d_1(u))^4 = 0 \end{aligned}$$

obtaining after some algebraic operations

$$\begin{aligned} d_1(u) &= u, \\ d_2(u) &= -\frac{\phi'''(1/M)}{4M^{1/2}\phi''(1/M)}\sum_{j=1}^{M}(a_j^t u)^2 a_j, \\ d_3(u) &= \frac{1}{96M}\left\{\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 \frac{15}{M}\left(\sum_{j=1}^{M} M(a_j^t u)^3 a_j - (u^t u)u\right)\right. \\ &\qquad \left. -\frac{7\phi^{IV}(1/M)}{\phi''(1/M)}\sum_{j=1}^{M}(a_j^t u)^3 a_j\right\}. \end{aligned}$$

The Jacobian of the transformation is given by

$$\begin{aligned} (\partial z/\partial u) &= I_r + n^{-1/2}\left(-\frac{\phi'''(1/M)}{2M^{1/2}\phi''(1/M)}\right)P_1 + n^{-1}\frac{1}{96M}\left\{\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2\frac{15}{M}\right. \\ &\qquad \left. \times\left(3MP_2 - 2u^t u - uu^t I_r\right) - 21\frac{\phi^{IV}(1/M)}{\phi''(1/M)}P_2\right\} + O(n^{-3/2}) \end{aligned}$$

where

$$P_1 = \sum_{j=1}^{M}(a_j^t u)a_j a_j^t \qquad \text{and} \qquad P_2 = \sum_{j=1}^{M}(a_j^t u)^2 a_j a_j^t.$$

In order to calculate the determinant of this matrix, the following general result is required where B and C are square $r \times r$ matrices and I_r is the $r \times r$ identity

$$\begin{aligned} &|I_r + n^{-1/2}B + n^{-1}C| \\ &= 1 + n^{-1/2}\sum_{i=1}^{r} b_{ii} + n^{-1}\left(\sum_{i=1}^{r} c_{ii} + \frac{1}{2}\sum_{i,j}^{r}(b_{ii}b_{jj} - b_{ij}b_{ji})\right) + O(n^{-3/2}). \end{aligned}$$

Hence it follows that

$$\begin{aligned} |\partial z/\partial u| &= 1 + n^{-1/2}\left(-\frac{\phi'''(1/M)}{2M^{1/2}\phi''(1/M)}\right)Q_1 + \frac{n^{-1}}{32M}\left((15MQ_2 - 5(2+r)u^t u\right. \\ &\qquad \left. +4MQ_1^2 - 4MQ_{12})\frac{1}{M}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 - 7\frac{\phi^{IV}(1/M)}{\phi''(1/M)}Q_2\right) + O(n^{-3/2}) \end{aligned}$$

where

$$Q_1 = \sum_{j=1}^{M}(a_j^t u)a_j^t a_j, \quad Q_2 = \sum_{j=1}^{M}(a_j^t u)^2 a_j^t a_j \quad \text{and} \quad Q_{12} = \sum_{k,j}^{M}(a_k^t u)(a_j^t u)(a_k^t a_j)^2.$$

Using the identities

$$Q_1 = \frac{T_1}{M^{1/2}}, \qquad Q_2 = \frac{T_2 - u^t u}{M} \qquad \text{and} \qquad Q_{12} = \frac{T_2 - 2u^t u}{M}$$

where z has been replaced by u on T_1 and T_2, we obtain that

$$|\partial z/\partial u| = 1 + n^{-1/2}\left(-\frac{\phi'''(1/M)}{2M\phi''(1/M)}\right)T_1 + \frac{n^{-1}}{32M^2}\left((15T_2 - 5(5+r)\,u^t u + 4T_1^2 - 4T_2 \right.$$
$$\left. + 8u^t u)\frac{1}{M}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 - 7\frac{\phi^{IV}(1/M)}{\phi''(1/M)}(T_2 - u^t u)\right) + O(n^{-3/2}). \quad (11)$$

Substituting u for z, it follows that

$$V_\phi(c) = |\Omega|^{1/2}\int_{u^t u < c}\cdots\int |\partial z/\partial u|\,\mathrm{d}u$$

i. e.

$$V_\phi(c) = |\Omega|^{1/2}\left\{M_r + n^{-1/2}\left(-\frac{\phi'''(1/M)}{2M\phi''(1/M)}\right)N_1\right.$$
$$+\frac{n^{-1}}{32M^2}\left\{(-(15+5r)N_2 + 11N_3 + 4N_4)\right.$$
$$\left.\left.\times\frac{1}{M}\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 - 7\frac{\phi^{IV}(1/M)}{\phi''(1/M)}(N_3 - N_2)\right\}\right\} + O(n^{-3/2})$$

where

$$M_r = \int_{u^t u < c}\cdots\int \mathrm{d}u = (\pi c)^{r/2}/\Gamma(1 + r/2),$$

$$N_1 = \sum_{j=1}^{M}\sqrt{M}\left(\sum_{k=1}^{r} a_{jk} I_k\right), \qquad N_2 = \sum_{k=1}^{r} I_{kk},$$

$$N_3 = \sum_{j=1}^{M} M\left(\sum_{k,m}^{r} a_{jk} a_{jm} I_{km}\right), \qquad N_4 = \sum_{i,j}^{M} M\left(\sum_{k,m}^{r} a_{ik} a_{jm} I_{km}\right),$$

and

$$I_k = \int_{u^t u < c}\cdots\int u_k\,\mathrm{d}u, \qquad I_{km} = \int_{u^t u < c}\cdots\int u_k u_m\,\mathrm{d}u.$$

Furthermore, from the proof of Theorem 2.1.8 of Read [18], we know that

$$I_k = 0,\ I_{km} = 0 \quad \text{for} \quad k \neq m \qquad \text{and} \qquad I_{kk} = \frac{M_r c}{r+2}, \quad k = 1, \ldots, r.$$

Hence

$$N_1 = 0, \qquad N_2 = (M-1)\frac{M_r c}{M+1},$$
$$N_3 = (M-1)M\frac{M_r c}{M+1}, \qquad N_4 = N_3 - (M-1)M\frac{M_r c}{M+1} = 0,$$

so

$$V_\phi(c) = \left(\frac{1}{M}\right)^{M/2} M_r \left\{1 + \frac{c(M-1)}{32M^2(M+1)n}\left(\left(\frac{\phi'''(1/M)}{\phi''(1/M)}\right)^2 (6(M-2)) \right.\right.$$
$$\left.\left. 7\frac{\phi^{IV}(1/M)}{\phi''(1/M)}(M-1)\right)\right\} + O(n^{-3/2}).$$

This is the result required and hence ends the proof. □

Remark 1. If we consider $\phi_\alpha(x) = (1-\alpha)^{-1}(x^\alpha - x)$ with $\alpha = 2$ then $\phi''(1/M) = -2$, $\phi'''(1/M) = 0$ and $\phi^{IV}(1/M) = 0$, and we obtain

$$J_1^{\phi_2} = P(\chi_r^2 < c) + \frac{(M-1)}{12n}\{P(\chi_r^2 < c)(-(M+1)) + 3MP(\chi_{r+2}^2 < c)$$
$$- 3(M-1)P(\chi_{r+4}^2 < c) + (M-2)P(\chi_{r+6}^2 < c) + O(n^{-3/2})$$

and

$$V_{\phi_2}(c) = \frac{(\pi c)^{r/2}}{\Gamma(1+r/2)}\left(\frac{1}{M}\right)^{M/2} + O(n^{-3/2})$$

that it is to say the result obtained by Yarnold [25].

The above approximation is closer to the exact distribution of the family $S_\phi(X/n, \pi_0)$ than the χ^2 approximation. However the effort required to calculate the second order approximation is substantial in comparison to calculating the χ^2 approximation.

Note that the J_1^ϕ term would be the Edgeworth expansion term if S_ϕ had a continuous distribution function. The term J_2^ϕ accounts for the error due to the discontinuity. Finally, the term $J_3^\phi = O(n^{-1})$ may be ignored as due to the asymptotic equivalence of the family S_ϕ discussed by Pardo [12], it follows that $n(J_3^{\phi_1} - J_3^{\phi_2}) \to 0$ as $n \to \infty$. Therefore any ϕ-dependent terms in J_3^ϕ will be $O(n^{-3/2})$. As in the expansion of S_ϕ distribution of Theorem 1 only includes terms larger than $O(n^{-3/2})$, J_3^ϕ can be viewed as independent of ϕ. Apart from this term can only cause a constant adjustment to the distribution function independent of ϕ, this evaluation is complex (see e. g. Yarnold [25] for the Pearson statistic).

4. NUMERICAL COMPUTATIONS

We compare the performances of the second order and chi-square approximations with the exact multinomial probabilities under equiprobable null hypothesis. We calculate the maximum approximation error incurred by $T_{\chi_{M-1}}$ and $T_S = J_1^\phi + \widehat{J}_2^\phi$ to T_E. The sign associated with the maximum difference is recorded to know if there has been an over-estimate or under-estimate. The computations are carried out using FORTRAN programs. The results are illustrated with the family S_{ϕ_α} for $\alpha \in (0,3]$; $M = 3,4,5,6$ and sample sizes $n = 10, 20$.

$M=3$	$n=10$		$n=20$	
α	T_χ	T_S	T_χ	T_S
.3	−.1879	−.0700	−.1218	−.0289
.5	−.1823	−.0776	−.1094	−.0218
.7	−.1679	−.0618	−.0982	−.0195
1	−.1542	−.0426	−.0935	−.0144
1.5	−.1406	−.0288	−.0866	−.0059
13/7	−.1280	−.0156	−.0822	−.0039
2	−.1162	−.0158	−.0916	−.0044
2.5	−.1397	.0170	−.1018	−.0138
3	−.1542	−.0419	−.1140	−.0208

$M=4$	$n=10$		$n=20$	
α	T_χ	T_S	T_χ	T_S
.3	−.2575	−.2193	−.1285	−.0546
.5	−.2364	−.2153	−.1165	−.0466
.7	−.2114	−.1926	−.1055	−.0344
1	−.1934	−.1217	−.0900	−.0178
1.5	−.1639	−.0265	−.0739	−.0055
13/7	−.1507	−.0207	−.0817	−.0069
2	−.1472	−.0313	−.0851	−.0074
2.5	−.1419	−.0357	−.0692	−.0153
3	−.1472	.0495	−.0718	−.0241

$M=5$	$n=10$		$n=20$	
α	T_χ	T_S	T_χ	T_S
.3	−.4257	−.4247	−.1638	−.0705
.5	−.4184	−.3947	−.1324	−.0570
.7	−.3515	−.2843	−.1206	−.0530
1	−.2281	−.1357	−.0784	−.0339
1.5	−.0830	−.0294	.0373	−.0048
13/7	.0752	.0283	.0337	.0064
2	.0762	.0378	−.0701	−.0084
2.5	.1013	.0697	.0601	.0194
3	.0885	.0990	.0401	.0437

$M=6$	$n=10$		$n=20$	
α	T_χ	T_S	T_χ	T_S
.3	−.7269	−.7164	−.1914	−.1520
.5	−.6855	−.6198	−.1713	−.1495
.7	−.5467	−.4045	−.1431	−.1277
1	−.3247	−.1671	−.1014	−.0683
1.5	−.1742	−.0560	−.0538	−.0117
13/7	−.1036	.0349	−.0444	.0103
2	−.1132	.0434	−.0568	.0127
2.5	−.0993	.0919	−.0293	.0322
3	−.1420	.1799	−.0673	.0716

The approximation T_S is the best since is the closest to the zero abscissa, i.e., the maximum approximation error resulting from using this approximation for the true distribution function T_E of S_{ϕ_α} is the closest to zero. There is more difference between T_S and $T_{\chi_{M-1}}$ for $M=3,4$ than $M=5,6$. Observe that we are comparing the T_S and $T_{\chi_{M-1}}$ asymptotic distributions of the chi-square statistic, S_{ϕ_2}. From this criterion we see that if we want to use the standard χ^2 approximation then we should use a α value in the range $[1.5,2]$.

(Received April 12, 1999.)

REFERENCES

[1] M. S. Ali and D. Silvey: A general class of coefficients of divergence of one distribution from another. J. Roy. Statist. Soc. Ser. B *28* (1966), 131–140.

[2] T. Bednarski and T. Ledwina: A note on a biasedness of tests of fit. Mathematische Operationsforschung und Statistik, Series Statistics *9* (1978), 191–193.

[3] J. Burbea: *J*-divergences and related topics. Encycl. 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] J. Burbea and C. R. Rao: On the convexity of higher order Jensen differences based on entropy function. IEEE Trans. Inform. Theory *28* (1982), 961–963.

[6] A. Cohen and H. B. Sackrowitz: Unbiasedness of the chi-square, likelihood ratio, and other goodness of fit tests for the equal cell case. Ann. Statist. *3* (1975), 959–964.

[7] N. Cressie and T. R. C. Read: Multinomial goodness of fit test. J. Roy. Statist. Soc. Ser. B *46* (1984), 440–464.

[8] I. Csiszár: Eine Informationtheoretische Ungleichung und ihre Anwendung auf den Bewis der Ergodizität von Markhoffschen Ketten. Publ. Math. Inst. Hungar. Acad. Sci. Ser. A *8* (1963), 85–108.

[9] P. E. Greenwood and M. S. Nikolin: A Guide to Chi-squared Testing. Wiley, New York 1996.

[10] F. Liese and I. Vajda: Convex Statistical Distances. Teubner, Leipzig 1987.

[11] H. B. Mann and A. Wald: On the choice of the number of intervals in the application of the chi-square test. Ann. Math. Statist. *13* (1942), 306–317.

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

[13] M. C. Pardo and I. Vajda: About distances of discrete distributions satisfying the data processing theorem of information theory. Trans. IEEE Inform. Theory *43* (1997), 4, 1288–1293.

[14] K. Pearson: On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophy Magazine *50* (1900), 157–172.

[15] C. R. Rao: Asymptotic efficiency and limiting information. In: Proc. 4th Berkeley Symp. on Math. Statist. Probab. *1*, Univ. of California Press, Berkeley 1961, pp. 531–546.

[16] C. R. Rao: Linear Statistical Inference and Its Applications. Second edition. Wiley, New York 1973.

[17] C. R. Rao: Analysis of Diversity: A Unified Approach. Technical Report No. 81-26, University of Pittsburhg 1981.

[18] T. R. C. Read: On Choosing a Goodness-of-fit test. Unpublished Ph.D. Thesis, Flinders University, South Australia 1982.

[19] T. R. C. Read: Closer asymptotic approximations for the distributions of the power divergence goodness-of-fit statistics. Ann. Inst. Statist. Math. *36* (1984), Part A, 59–69.

[20] T. R. C. Read and N. Cressie: Goodness of Fit Statistics for Discrete Multivariate Data. Springer, New York 1988.

[21] B. Schorr: On the choice of the class intervals in the application of the chi-square test. Math. Operationsforsch. Statist. *5* (1974), 357–377.

[22] M. Siotani and Y. Fujikoshi: Asymptotic approximations for the distributions of multinomial goodness-of-fit statistics. Hiroshima Math. J. *14* (1984), 115–124.

[23] H. A. Sturges: The choice of a class interval. J. Amer. Statist. Assoc. *21* (1926), 65–66.

[24] I. Vajda: Theory of Statistical Inference and Information. Kluwer Academic Publishers, Dordrecht – Boston 1989.

[25] J. K. Yarnold: Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set. Ann. of Math. Statist. *43* (1972), 1566–1580.

[26] 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. María del Carmen Pardo, Departamento de Estadística e I.O., Facultad de Matemáticas, Universidad Complutense de Madrid, 28040-Madrid. Spain.
e-mail: mcpardo@eucmax.sim.ucm.es