AN ASYMPTOTIC APPROXIMATION FOR THE DISTRIBUTION OF ϕ -DIVERGENCE MULTINOMIAL GOODNESS-OF-FIT STATISTIC UNDER LOCAL ALTERNATIVES

Nobuhiro Taneichi*, Yuri Sekiya** and Akio Suzukawa*

On the goodness-of-fit test for multinomial distribution, Zografos *et al.* (1990) proposed the ϕ -divergence family of statistics, which includes the power divergence family of statistics as a special case. They showed that under null hypothesis, the members of the ϕ -divergence family of statistics all have an asymptotically equivalent chi-square distribution. Furthermore, Menéndez *et al.* (1997) derived an asymptotic expansion for the null distribution of the ϕ -divergence statistic. In this paper, we derive an approximation for the distribution of the ϕ -divergence statistic under local alternatives. The approximation is based on the continuous term of the asymptotic expansion for the distribution of the ϕ -divergence statistic. By using the approximation, we propose a new approximation for the power of the statistic. The results are generalizations of those derived by Taneichi *et al.* which discussed the power divergence statistic. We numerically investigate the accuracy of the approximation when two types of concrete ϕ -divergence statistics are applied. By the numerical investigation, we find that the present approximation performs better than the other approximations.

Key words and phrases: ϕ -divergence statistic, Goodness-of-fit, Multinomial distribution.

1. Introduction

A measure for the dissimilarity between distributions, which is called ϕ -divergence, was proposed independently by Csiszár (1967) and Ali and Silvey (1966). The ϕ -divergence measure between discrete distributions $\mathbf{p} = (p_1, \dots, p_k)'$ and $\mathbf{q} = (q_1, \dots, q_k)'$ is defined by

$$(1.1) \quad D_\phi(\mathbf{p}, \mathbf{q}) = \sum_{j=1}^k q_j \phi\left(\frac{p_j}{q_j}\right),$$

where $\phi(t)$ is a real convex function for $t > 0$ with

$$\begin{aligned} \phi(0) &= \lim_{t \rightarrow +0} \phi(t), \\ 0\phi\left(\frac{x}{0}\right) &= \lim_{y \rightarrow +0} y\phi\left(\frac{x}{y}\right) = x \lim_{t \rightarrow +0} t\phi\left(\frac{1}{t}\right), \quad \text{if } x > 0, \end{aligned}$$

Received September 20, 2000. Revised May 25, 2001. Accepted June 4, 2001.

*Obihiro University of Agriculture and Veterinary Medicine, Inada-chō Nishi 2 Sen, Obihiro 080-8555, Japan.

**Kushiro Campus, Hokkaido University of Education, Shiroyama 1-15-55, Kushiro 085-8580, Japan.

and $0\phi(0/0) = 0$.

Let $\mathbf{X} = (X_1, \dots, X_k)'$ be distributed according to a multinomial distribution $\text{Mult}_k(n, \boldsymbol{\pi})$, where $\sum_{j=1}^k X_j = n$, $\sum_{j=1}^k \pi_j = 1$, $0 < \pi_j < 1$, $(j = 1, \dots, k)$ and $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)'$ is an unknown probability vector. In order to test the null hypothesis

$$(1.2) \quad H_0 : \boldsymbol{\pi} = \mathbf{p}$$

for some completely specified probability vector $\mathbf{p} = (p_1, \dots, p_k)'$, Zografos *et al.* (1990) introduced the ϕ -divergence family of statistics

$$C_\phi = \frac{2n}{\phi''(1)} \left\{ D_\phi \left(\frac{\mathbf{X}}{n}, \mathbf{p} \right) - \phi(1) \right\}, \quad (\phi''(1) \neq 0).$$

Without loss of generality, we consider a twice continuously differentiable convex function $\phi(t)$ for $t > 0$, satisfying $\phi(1) = \phi'(1) = 0$ and $\phi''(1) = 1$ (Pardo *et al.* (1999)). So, the ϕ -divergence family of statistics is defined by

$$C_\phi = 2n D_\phi \left(\frac{\mathbf{X}}{n}, \mathbf{p} \right).$$

When we choose the following convex function ϕ_a^P as ϕ , ϕ -divergence measure becomes the power divergence measure (Read and Cressie (1988)).

$$\phi_a^P(t) = \begin{cases} \{a(a+1)\}^{-1} \{t^{a+1} - t + a(1-t)\} & (a \neq 0, -1) \\ t \ln t + 1 - t & (a = 0) \\ -\ln t - 1 + t & (a = -1). \end{cases}$$

Therefore, the ϕ -divergence family of statistics C_ϕ includes the power divergence family of statistics proposed by Cressie and Read (Cressie and Read (1984) and Read and Cressie (1988)) as $C_{\phi_a^P}$. It is well known that the family of the statistics $C_{\phi_a^P}$ includes Pearson's X^2 statistic ($a = 1$), the loglikelihood ratio statistic ($a = 0$), the Freeman-Tukey statistic ($a = -1/2$), the modified loglikelihood ratio statistic ($a = -1$), the Neyman modified X^2 statistic ($a = -2$), and the statistic recommended by Cressie and Read ($a = 2/3$). The power divergence is used not only for a test statistic, but also for representation of the degree of departure from symmetry for square contingency tables (Tomizawa *et al.* (1998)). On the other hand, Rukhin (1994) considered the family of statistics given by the following convex function ϕ_a^R as ϕ .

$$\phi_a^R(t) = \frac{(t-1)^2}{2(a+t(1-a))}, \quad a \in [0, 1].$$

Then the statistic based on ϕ_a^R is defined as

$$(1.3) \quad C_{\phi_a^R} = n \sum_{j=1}^k \frac{(\hat{p}_j - p_j)^2}{\{ap_j + (1-a)\hat{p}_j\}}, \quad a \in [0, 1],$$

where $\hat{p}_j = X_j/n$, ($j = 1, \dots, k$). The family of the statistics $C_{\phi_a^L}$ includes Pearson's X^2 statistic ($a = 1$) and the Neyman modified X^2 statistic ($a = 0$). Furthermore, Lin (1991) considered the family of divergences given by the following convex function ϕ_a^L as ϕ .

$$\phi_a^L(t) = \{a(1-a)\}^{-1} \{at \ln t - (at + 1 - a) \ln(at + 1 - a)\}, \quad a \in [0, 1].$$

Then the statistic based on ϕ_a^L is defined as

$$(1.4) \quad C_{\phi_a^L} = \frac{2n}{a(1-a)} \left\{ - \sum_{j=1}^k s_j \ln s_j + (1-a) \sum_{j=1}^k p_j \ln p_j + a \sum_{j=1}^k \hat{p}_j \ln \hat{p}_j \right\},$$

$$a \in [0, 1],$$

where $s_j = a\hat{p}_j + (1-a)p_j$, ($j = 1, \dots, k$). The family of the statistics $C_{\phi_a^L}$ includes the loglikelihood ratio statistic ($a = 0$) and the modified loglikelihood ratio statistic ($a = 1$) by considering continuity in a . The divergence measure $D_{\phi_a^L}$ becomes J-divergence proposed by Burbea and Rao (1982). Therefore, the statistic $C_{\phi_a^L}$ coincides with the statistic J_1^a based on J-divergence studied by Sekiya (1990). Taneichi (1993) derived the asymptotic distribution of the statistic $C_{\phi_{0.5}^L}$ (or $J_1^{0.5}$) under a model containing unspecified parameters. Furthermore, Taneichi and Sato (1989) and Taneichi (1995) discussed the asymptotic distribution of the minimum J-divergence estimator.

Zografos *et al.* (1990) showed that under the null hypothesis H_0 , all of the members of the ϕ -divergence family of the statistics C_ϕ are asymptotically distributed according to the central chi-square distribution with $k - 1$ degrees of freedom. If we consider further asymptotic approximations for the distribution of C_ϕ , it is necessary to refer to the following studies. By using the Euler-Mclaurin summation formula, Esséen (1945) derived the asymptotic expansion for the distribution function of the sum of i.i.d. lattice-valued random variables. By using the multidimensional version of the Euler-Mclaurin summation formula, Ranga Rao (1961) derived the asymptotic expansion for the probability that the normalized sum of i.i.d. lattice-valued random vectors is contained in a Borel set B . (Details are shown in Bhattacharya and Ranga Rao (1976).) Using Ranga Rao's expansion, Yarnold (1972) obtained a somewhat simple expansion when B is a convex set which includes a term for the number of lattice points inside the region of interest. Yarnold (1972) also applied it to the null distribution of Pearson's X^2 statistic and derived an asymptotic expansion for the null distribution of X^2 . In a similar fashion to X^2 , asymptotic expansions for the null distributions of the loglikelihood ratio statistic and the Freeman-Tukey statistic were obtained by Siotani and Fujikoshi (1984), that of the power divergence statistic was obtained by Read (1984), and that of C_ϕ was obtained by Menéndez *et al.* (1997). On the other hand, omitting the discontinuous term, Taneichi *et al.* derived a formal asymptotic expansion for the distribution of the power divergence statistic under local alternative hypotheses

$$(1.5) \quad H_{1,n} : \pi_j = p_j + \frac{c_j}{\sqrt{n}} \quad \text{for all } j = 1, \dots, k, \quad \sum_{j=1}^k c_j = 0.$$

In this paper, by generalizing the results of Taneichi *et al.*, we consider an asymptotic approximation for the distribution of the ϕ -divergence statistic C_ϕ under local alternatives $H_{1,n}$. As a special case, the local alternative hypotheses $H_{1,n}$ coincide with the null hypothesis H_0 when $c_j = 0$ ($j = 1, \dots, k$). So, in Section 2, the results of the asymptotic expansions for the null distributions of the multinomial goodness-of-fit test statistics are summarized. In Section 3, omitting the discontinuous term, we derive a formal asymptotic expansion for the distribution of C_ϕ under local alternatives and propose a new approximation for the power of the statistic. In Section 4, accuracy of the approximation is numerically investigated when we apply the approximation to the statistic $C_{\phi_a^R}$ given by (1.3) and the statistic $C_{\phi_a^L}$ given by (1.4).

2. Asymptotic expansions for the distributions of the goodness-of-fit test statistics under H_0

In this section, we summarize the results of the asymptotic expansions for the null distributions of the multinomial goodness-of-fit test statistics. Throughout this section let $\mathbf{X} = (X_1, \dots, X_k)'$ be distributed according to $\text{Mult}_k(n, \mathbf{p})$. We put $r = k - 1$. Let

$$U_j = \frac{X_j - np_j}{\sqrt{n}}, \quad (j = 1, \dots, k)$$

and $\mathbf{U} = (U_1, \dots, U_r)'$. Then $\mathbf{U}$ is a lattice random vector which takes values in the set

$$(2.1) \quad L = \left\{ \mathbf{u} = (u_1, \dots, u_r)' : \mathbf{u} = \frac{\tilde{\mathbf{x}} - n\tilde{\mathbf{p}}}{\sqrt{n}}, \tilde{\mathbf{x}} \in M \right\},$$

where

$$(2.2) \quad M = \left\{ \tilde{\mathbf{x}} = (x_1, \dots, x_r)' : \right. \\ \left. x_1, \dots, x_r \text{ are non-negative integers and } \sum_{j=1}^r x_j \leq n \right\}$$

and $\tilde{\mathbf{p}} = (p_1, \dots, p_r)'$. We need the following lemma, which states a local Edgeworth expansion under H_0 given by (1.2).

LEMMA 2.1. Let $\mathbf{u} = (\tilde{\mathbf{x}} - n\tilde{\mathbf{p}})/\sqrt{n}$. Then for any $\tilde{\mathbf{x}} \in M$,

$$\Pr\{\mathbf{U} = \mathbf{u} \mid H_0\} = n^{-r/2} f_0(\mathbf{u}) \left\{ 1 + \frac{1}{\sqrt{n}} h_1(\mathbf{u}) + \frac{1}{n} h_2(\mathbf{u}) + O(n^{-3/2}) \right\},$$

where

$$f_0(\mathbf{u}) = (2\pi)^{-r/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{u}' \Omega^{-1} \mathbf{u} \right\},$$

$$\Omega = \text{diag}(p_1, \dots, p_r) - \tilde{\mathbf{p}}\tilde{\mathbf{p}}',$$

$$h_1(\mathbf{u}) = -\frac{1}{2} \sum_{j=1}^r \frac{u_j}{p_j} + \frac{1}{6} \sum_{j=1}^r \frac{u_j^3}{p_j^3},$$

$$h_2(\mathbf{u}) = \frac{1}{2}\{h_1(\mathbf{u})\}^2 + \frac{1}{12}(1 - S_1^0) + \frac{1}{4}\sum_{j=1}^k \frac{u_j^2}{p_j^2} - \frac{1}{12}\sum_{j=1}^k \frac{u_j^4}{p_j^3},$$

$$u_k = -\sum_{j=1}^r u_j,$$

and

$$S_1^0 = \sum_{j=1}^k \frac{1}{p_j}.$$

By using Theorem 22.1 of Bhattacharya and Ranga Rao (1976), this lemma was obtained by Siotani and Fujikoshi (1984). Yarnold (1972) gave an asymptotic expansion for $\Pr\{\mathbf{U} \in G \mid H_0\}$ when G is an extended convex set. A set G is called an extended convex set if G has the representation

$$G = \{\mathbf{u} = (u_1, \dots, u_r)' : \alpha_j(\mathbf{u}_j^*) < u_j < \beta_j(\mathbf{u}_j^*), \\ \mathbf{u}_j^* = (u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_r)' \in G_j\}$$

for every $j \in \{1, 2, \dots, r\}$, where $G_j \subset \mathbf{R}^{r-1}$ and α_j, β_j are continuous functions on $\mathbf{R}^{r-1}$ (Yarnold (1972)). By using the notation of Siotani and Fujikoshi (1984), the Yarnold's result is described as follows:

THEOREM 2.1. *If G is a bounded extended convex set, then*

$$(2.3) \quad \Pr\{\mathbf{U} \in G \mid H_0\} = K_1 + K_2 + K_3 + O(n^{-3/2}),$$

where

$$K_1 = \int \cdots \int_G f_0(\mathbf{u}) \left\{ 1 + \frac{1}{\sqrt{n}} h_1(\mathbf{u}) + \frac{1}{n} h_2(\mathbf{u}) \right\} du_1 \cdots du_r,$$

$$K_2 = -\frac{1}{\sqrt{n}} \sum_{j=1}^r n^{-(r-j)/2} \sum_{u_{j+1} \in L_{j+1}} \cdots \sum_{u_r \in L_r} \\ \cdot \int \cdots \int_{G_j} [\lambda_1(\sqrt{n}u_j + np_j) f_0(\mathbf{u})]_{\alpha_j(\mathbf{u}_j^*)}^{\beta_j(\mathbf{u}_j^*)} du_1 \cdots du_{j-1},$$

$$K_3 = \frac{1}{n} \sum_{j=1}^r n^{-(r-j)/2} \sum_{u_{j+1} \in L_{j+1}} \cdots \sum_{u_r \in L_r} \\ \cdot \int \cdots \int_{G_j} \left[-\lambda_1(\sqrt{n}u_j + np_j) h_1(\mathbf{u}) f_0(\mathbf{u}) \right. \\ \left. + \lambda_2(\sqrt{n}u_j + np_j) \frac{\partial}{\partial u_j} f_0(\mathbf{u}) \right]_{\alpha_j(\mathbf{u}_j^*)}^{\beta_j(\mathbf{u}_j^*)} du_1 \cdots du_{j-1},$$

$$\lambda_1(t) = t - [t] - \frac{1}{2},$$

$\lambda_2(t)$ is the real-valued periodic function with period one such that

$$\lambda_2(t) = \frac{1}{2} \left(t^2 - t + \frac{1}{6} \right) \quad \text{for } t \in [0, 1],$$

$$\begin{aligned}
[h(\mathbf{u})]_{\alpha_j(\mathbf{u}_j^*)}^{\beta_j(\mathbf{u}_j^*)} &= h(u_1, \dots, u_{j-1}, \beta_j(\mathbf{u}_j^*), u_{j+1}, \dots, u_r) \\
&\quad - h(u_1, \dots, u_{j-1}, \alpha_j(\mathbf{u}_j^*), u_{j+1}, \dots, u_r), \\
L_j &= \left\{ u_j : u_j = \frac{x_j - np_j}{\sqrt{n}} \text{ and } x_j \text{ is integer} \right\},
\end{aligned}$$

and $f_0(\cdot)$ and $h_j(\cdot)$, ($j = 1, 2$) are given in Lemma 2.1. Furthermore, it holds that $K_2 = O(n^{-1/2})$ and $K_3 = O(n^{-1})$.

The K_1 term can be regarded as the Edgeworth expansion for a continuous distribution. The K_2 term is a discontinuous term to account for the discontinuity in $\mathbf{U}$. Using Lemma 2.1 and Theorem 2.1 we can obtain an asymptotic expansion of the null distribution of C_ϕ . Let

$$(2.4) \quad G_\eta^\phi = \{\mathbf{u} = (u_1, \dots, u_r)' : C_\phi(\mathbf{u}) < \eta\},$$

where

$$C_\phi(\mathbf{u}) = 2n \sum_{j=1}^k p_j \phi \left(1 + \frac{u_j}{\sqrt{np_j}} \right).$$

Since the set G_η^ϕ is a bounded extended convex set, we can write $\Pr\{C_\phi(\mathbf{U}) < \eta \mid H_0\} = \Pr\{\mathbf{U} \in G_\eta^\phi \mid H_0\}$ as the formula (2.3) with $G = G_\eta^\phi$.

THEOREM 2.2. *If ϕ is forth times differentiable and $\phi^{(4)}(t)$ is continuous at $t = 1$, then K_1 and K_2 in the case of $G = G_\eta^\phi$ are evaluated as follows:*

$$(2.5) \quad K_1 = \Pr\{\chi_r^2 < \eta\} + \frac{1}{24n} \sum_{j=0}^3 r'_j \Pr\{\chi_{r+2j}^2 < \eta\} + o(n^{-1}),$$

where

$$\begin{aligned}
r'_0 &= 2(1 - S_1^0), \\
r'_1 &= 3(-k^2 - 2k + 3S_1^0) + 6T^3(-k^2 + S_1^0) \\
&\quad + 3T^4(2k - 1 - S_1^0) + (T^3)^2(-3k^2 - 6k + 4 + 5S_1^0), \\
r'_2 &= 6(k^2 + 2k - 1 - 2S_1^0) + 4T^3(3k^2 + 3k - 2 - 4S_1^0) \\
&\quad + 3T^4(-2k + 1 + S_1^0) + 2(T^3)^2(3k^2 + 6k - 4 - 5S_1^0), \\
r'_3 &= (1 + T^3)^2(-3k^2 - 6k + 4 + 5S_1^0), \\
T^l &= \phi^{(l)}(1),
\end{aligned}$$

S_1^0 is defined in Lemma 2.1, and χ_r^2 denotes a central chi-square random variable with r degrees of freedom. K_2 can be approximated to first order by

$$(2.6) \quad \hat{K}_2 = \frac{\{N(\eta) - n^{r/2}V(\eta)\}e^{-\eta/2}}{\{(2\pi n)^r \prod_{j=1}^k p_j\}^{1/2}},$$

where

$$V(\eta) = V_0 \left[1 + \frac{\eta \{ (T^3)^2 (5S_1^0 - 3k^2 - 6k + 4) - 3T^4 (S_1^0 - 2k + 1) \}}{24n(k+1)} \right] \\ + O(n^{-3/2}), \\ V_0 = \frac{\{ (\pi\eta)^r \prod_{j=1}^k p_j \}^{1/2}}{\Gamma\left(\frac{k+1}{2}\right)},$$

$N(\eta)$ is the number of members in $\{\mathbf{u} : \mathbf{u} \in L, \mathbf{u} \in G_\eta^\phi\}$, and $\Gamma(\cdot)$ is the gamma function.

The expansion (2.5) and $\hat{K}_2$ given by (2.6) were derived in Menéndez *et al.* (1997). However, the necessary conditions for Theorem 2.2 about the function ϕ were not shown in their paper. Although the calculus of K_3 in the case of $G = G_\eta^\phi$ is too complicated, we know by Theorem 2.1 that $K_3 = O(n^{-1})$. Therefore, we obtain

$$K'_1 + \hat{K}_2,$$

where

$$K'_1 = \Pr\{\chi_r^2 < \eta\} + \frac{1}{24n} \sum_{j=0}^3 r'_j \Pr\{\chi_{r+2j}^2 < \eta\},$$

as an asymptotic approximation to $\Pr\{C_\phi(\mathbf{U}) < \eta \mid H_0\}$. The approximation $\hat{K}_2$ does not look so complicated. However, it is very hard to calculate the value of $\hat{K}_2$ when n and k are not small, because of the difficulty for counting $N(\eta)$ in (2.6). So, the use of the approximation is not so practical, except when n and k are small.

3. An asymptotic approximation for the distribution of ϕ -divergence statistic under $H_{1,n}$

In this section, we derive an approximation for the distribution of the ϕ -divergence statistic C_ϕ under local alternatives (1.5). So, throughout this section let $\mathbf{X} = (X_1, \dots, X_k)'$ be distributed according to $\text{Mult}_k(n, \mathbf{q})$, where $\mathbf{q} = \mathbf{p} + \mathbf{c}/\sqrt{n}$ and $\mathbf{c} = (c_1, \dots, c_k)'$. Let

$$Y_j = \frac{X_j - np_j}{\sqrt{n}}, \quad (j = 1, \dots, k)$$

and $\mathbf{Y} = (Y_1, \dots, Y_k)'$. Then $\mathbf{Y}$ is a lattice random vector which takes values in the set L defined by (2.1). A local Edgeworth expansion under $H_{1,n}$ is given in the following lemma.

LEMMA 3.1. *For each $\tilde{\mathbf{x}} \in M$, where M is defined by (2.2), let $\mathbf{y} = (\tilde{\mathbf{x}} - n\tilde{\mathbf{p}})/\sqrt{n}$. Then*

$$\Pr\{\mathbf{Y} = \mathbf{y} \mid H_{1,n}\} = n^{-r/2} f(\mathbf{y}) \left\{ 1 + \frac{1}{\sqrt{n}} g_1(\mathbf{y}) + \frac{1}{n} g_2(\mathbf{y}) + o(n^{-1}) \right\},$$

where

$$\begin{aligned}
 f(\mathbf{y}) &= (2\pi)^{-r/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2}(\mathbf{y} - \tilde{\mathbf{c}})' \Omega^{-1} (\mathbf{y} - \tilde{\mathbf{c}}) \right\}, \\
 \tilde{\mathbf{c}} &= (c_1, \dots, c_r)', \\
 g_1(\mathbf{y}) &= -\frac{1}{2} \sum_{j=1}^k \frac{y_j}{p_j} + \frac{1}{6} \sum_{j=1}^k \frac{y_j^3}{p_j^2} - \frac{1}{2} \sum_{j=1}^k \frac{c_j^2 y_j}{p_j^2} + \frac{1}{3} S_2^3, \\
 g_2(\mathbf{y}) &= \frac{1}{2} \{g_1(\mathbf{y})\}^2 + \frac{1}{12} (1 - S_1^0) + \frac{1}{4} \sum_{j=1}^k \frac{y_j^2}{p_j^2} - \frac{1}{12} \sum_{j=1}^k \frac{y_j^4}{p_j^3} + \frac{1}{3} \sum_{j=1}^k \frac{c_j^3 y_j}{p_j^3} - \frac{1}{4} S_3^4, \\
 y_k &= -\sum_{j=1}^r y_j, \quad c_k = -\sum_{j=1}^r c_j, \\
 S_m^l &= \sum_{j=1}^k \frac{(c_j)^l}{(p_j)^m},
 \end{aligned}$$

and Ω is defined in Lemma 2.1.

In a similar fashion to the proof of Theorem 22.1 of Bhattacharya and Ranga Rao (1976), this lemma was proved by Taneichi *et al.*. As a special case, the results of Lemma 3.1 coincide with the results of Lemma 2.1 when $\mathbf{c} = \mathbf{0}$. The distribution of the ϕ -divergence statistic C_ϕ under $H_{1,n}$ can be expressed as

$$\Pr\{C_\phi(\mathbf{Y}) < \eta \mid H_{1,n}\} = \Pr\{\mathbf{Y} \in G_\eta^\phi \mid H_{1,n}\},$$

where G_η^ϕ is defined by (2.4). By almost the same arguments as in Taneichi *et al.*, which discussed the power divergence statistic, we find the following results. The discontinuous term for the distribution of C_ϕ under $H_{1,n}$ which corresponds to K_2 given in (2.3) in the null hypothesis case is expressed in a very complicated form. So, under $H_{1,n}$, it does not have a simple asymptotic approximation such as $\hat{K}_2$ given by (2.6) in the null hypothesis case. Therefore, omitting the discontinuous term, we consider the following approximation for the distribution of C_ϕ under $H_{1,n}$.

$$\Pr\{C_\phi(\mathbf{Y}) < \eta \mid H_{1,n}\} \approx J,$$

where

$$J = \int \cdots \int_{G_\eta^\phi} f(\mathbf{y}) \left\{ 1 + \frac{1}{\sqrt{n}} g_1(\mathbf{y}) + \frac{1}{n} g_2(\mathbf{y}) \right\} dy_1 \cdots dy_r$$

is the multivariate Edgeworth expansion for a continuous distribution.

THEOREM 3.1. *If ϕ is forth times differentiable and $\phi^{(4)}(t)$ is continuous at $t = 1$, then J is evaluated as*

$$\begin{aligned}
 (3.1) \quad J &= \Pr\{\chi_r^2(\delta) < \eta\} + \frac{1}{6\sqrt{n}} \sum_{j=0}^3 w_j \Pr\{\chi_{r+2j}^2(\delta) < \eta\} \\
 &\quad + \frac{1}{72n} \sum_{j=0}^6 r_j \Pr\{\chi_{r+2j}^2(\delta) < \eta\} + o(n^{-1}),
 \end{aligned}$$

where

$$\begin{aligned}
w_0 &= 2S_2^3, \\
w_1 &= -3\{(S_1^1 + S_2^3) + T^3 S_1^1\}, \\
w_2 &= 3S_1^1 + T^3(3S_1^1 - S_2^3), \\
w_3 &= (1 + T^3)S_2^3, \\
r_0 &= 3r'_0 - 18S_3^4 + 4(S_2^3)^2, \\
r_1 &= 3r'_1 + 3\{-6kS_1^2 + 6S_2^2 + 11S_3^4 - 3(S_1^2)^2 - 4(S_2^3)^2 - 4S_1^1 S_2^3\} \\
&\quad + 6T^3\{-3kS_1^2 + 3S_2^2 - 2S_1^1 S_2^3\}, \\
r_2 &= 3r'_2 + 3\{6kS_1^2 + 3(S_1^1)^2 + 3(S_2^3)^2 + 10S_1^1 S_2^3\} \\
&\quad + 2T^3\{9S_3^4 + 9(S_1^1)^2 - 9(S_1^2)^2 - 2(S_2^3)^2 + 15S_1^1 S_2^3\} + 18T^4\{S_1^2 - S_2^2\} \\
&\quad + 9(T^3)^2\{-2(k+2)S_1^2 + 4S_2^2 + (S_1^1)^2\}, \\
r_3 &= 3r'_3 + 2\{9(k+2)S_1^2 - 27S_2^2 - 9S_3^4 - 9(S_1^1)^2 + 9(S_1^2)^2 + 2(S_2^3)^2 - 9S_1^1 S_2^3\} \\
&\quad + 2T^3\{9(3k+4)S_1^2 - 45S_2^2 - 9S_3^4 - 18(S_1^1)^2 + 9(S_1^2)^2 + 5(S_2^3)^2 - 6S_1^1 S_2^3\} \\
&\quad + 3T^4\{-6S_1^2 + 6S_2^2 - S_3^4\} \\
&\quad + 3(T^3)^2\{12(k+2)S_1^2 - 24S_2^2 + 3S_3^4 - 6(S_1^1)^2 - 3(S_1^2)^2 + 2S_1^1 S_2^3\}, \\
r_4 &= 3\{-6(k+2)S_1^2 + 12S_2^2 - 2S_3^4 + 3(S_1^1)^2 - 2(S_2^3)^2 - 2S_1^1 S_2^3\} \\
&\quad + 6T^3\{-6(k+2)S_1^2 + 12S_2^2 - 3S_3^4 + 3(S_1^1)^2 + 3(S_1^2)^2 - (S_2^3)^2 - 3S_1^1 S_2^3\} \\
&\quad + 3T^4 S_3^4 + (T^3)^2 \\
&\quad \cdot \{-18(k+2)S_1^2 + 36S_2^2 - 18S_3^4 + 9(S_1^1)^2 + 18(S_1^2)^2 + (S_2^3)^2 - 12S_1^1 S_2^3\}, \\
r_5 &= 3\{3S_3^4 - 3(S_1^1)^2 + 2S_1^1 S_2^3\} + 2T^3\{9S_3^4 - 9(S_1^1)^2 - (S_2^3)^2 + 6S_1^1 S_2^3\} \\
&\quad + (T^3)^2\{9S_3^4 - 9(S_1^1)^2 - 2(S_2^3)^2 + 6S_1^1 S_2^3\}, \\
r_6 &= (1 + T^3)^2(S_2^3)^2,
\end{aligned}$$

S_m^l is defined in Lemma 3.1, r'_j and T^l are defined in Theorem 2.2, and $\chi_r^2(\delta)$ denotes a noncentral chi-square random variable with r degrees of freedom and noncentrality parameter $\delta = S_1^2$.

PROOF. If we regard

$$f(\mathbf{y}) \left\{ 1 + \frac{1}{\sqrt{n}} g_1(\mathbf{y}) + \frac{1}{n} g_2(\mathbf{y}) \right\}$$

as the continuous density function of $\mathbf{Y}$, then we can regard J as the distribution function of $C_\phi(\mathbf{Y})$. In order to evaluate J , we calculate the characteristic function of $C_\phi(\mathbf{Y})$ as follows:

$$\begin{aligned}
(3.2) \quad \psi(s) &= \int \cdots \int_{R^r} \exp \{is C_\phi(\mathbf{y})\} f(\mathbf{y}) \\
&\quad \cdot \left\{ 1 + \frac{1}{\sqrt{n}} g_1(\mathbf{y}) + \frac{1}{n} g_2(\mathbf{y}) \right\} dy_1 \cdots dy_r.
\end{aligned}$$

From the assumption that ϕ is forth times differentiable and $\phi^{(4)}(t)$ is continuous at $t = 1$, we can expand $\phi(1 + y_j/\sqrt{np_j})$ as

$$\phi\left(1 + \frac{y_j}{\sqrt{np_j}}\right) = \frac{T^1 y_j}{\sqrt{np_j}} + \frac{y_j^2}{2np_j^2} + \frac{T^3 y_j^3}{6n\sqrt{np_j^3}} + \frac{T^4 y_j^4}{24n^2 p_j^4} + o(n^{-2})$$

for large n and fixed y_j . Therefore $C_\phi(\mathbf{y})$ can be expanded as

$$(3.3) \quad C_\phi(\mathbf{y}) = \mathbf{y}'\Omega^{-1}\mathbf{y} + \frac{T^3}{3\sqrt{n}} \sum_{j=1}^k \frac{y_j^3}{p_j^2} + \frac{T^4}{12n} \sum_{j=1}^k \frac{y_j^4}{p_j^3} + o(n^{-1})$$

for large n and fixed $\mathbf{y}$. Substituting (3.3) into (3.2), we obtain

$$\psi(s) = \int \cdots \int_{R^r} f(\mathbf{y}) \exp\{is \mathbf{y}'\Omega^{-1}\mathbf{y}\} A(\mathbf{y}) dy_1 \cdots dy_r + o(n^{-1}),$$

where

$$A(\mathbf{y}) = 1 + \frac{1}{\sqrt{n}} \left\{ \frac{isT^3}{3} \sum_{j=1}^k \frac{y_j^3}{p_j^2} + g_1(\mathbf{y}) \right\} + \frac{1}{n} \left\{ \frac{isT^4}{12} \sum_{j=1}^k \frac{y_j^4}{p_j^3} + \frac{(isT^3)^2}{18} \left(\sum_{j=1}^k \frac{y_j^3}{p_j^2} \right)^2 + \frac{isT^3}{3} g_1(\mathbf{y}) \sum_{j=1}^k \frac{y_j^3}{p_j^2} + g_2(\mathbf{y}) \right\}.$$

Let $\sigma^2 = (1 - 2is)^{-1}$ and

$$f^*(\mathbf{y}) = (2\pi)^{-r/2} \sigma^{-r} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y} - \sigma^2 \mathbf{c})' \Omega^{-1} (\mathbf{y} - \sigma^2 \mathbf{c}) \right\}.$$

Then $f^*(\mathbf{y})$ is the density function of the r -variate normal distribution $N_r(\sigma^2 \mathbf{c}, \sigma^2 \Omega)$, and $\psi(s)$ can be expressed as follows:

$$\psi(s) = \sigma^r \exp(is\sigma^2 \mathbf{c}' \Omega^{-1} \mathbf{c}) \int \cdots \int_{R^r} f^*(\mathbf{y}) A(\mathbf{y}) dy_1 \cdots dy_r + o(n^{-1}).$$

Using the moment formulae for the multivariate normal variate, we obtain

$$(3.4) \quad \psi(s) = (1 - 2is)^{-r/2} \exp\left(\frac{is}{1 - 2is} \delta\right) \times \left[1 + \frac{1}{6\sqrt{n}} \sum_{j=0}^3 w_j (1 - 2is)^{-j} + \frac{1}{72n} \sum_{j=0}^6 r_j (1 - 2is)^{-j} \right] + o(n^{-1}).$$

Inverting (3.4) we obtain (3.1). We have completed the proof of Theorem 3.1. $\square$

In particular, we know from Theorem 3.1 that under $H_{1,n}$, C_ϕ is asymptotically distributed according to the noncentral chi-square distribution with $k - 1$ degrees of freedom and noncentrality parameter δ . As a special case, the results

of Theorem 3.1 coincide with the first half of the results of Theorem 2.2 when $\mathbf{c} = \mathbf{0}$. On the basis of Theorem 3.1, we have an approximate expression

$$\begin{aligned} \Pr\{C_\phi(\mathbf{Y}) < \eta \mid H_{1,n}\} &\approx \Pr\{\chi_r^2(\delta) < \eta\} + \frac{1}{6\sqrt{n}} \sum_{j=0}^3 w_j \Pr\{\chi_{r+2j}^2(\delta) < \eta\} \\ &\quad + \frac{1}{72n} \sum_{j=0}^6 r_j \Pr\{\chi_{r+2j}^2(\delta) < \eta\} \end{aligned}$$

for large n . When n is large, therefore, we can approximate the power of C_ϕ against $H_{1,n}$ as

$$\begin{aligned} (3.5) \quad 1 - \Pr\{\chi_r^2(\delta) < c_0^\phi(\alpha)\} &- \frac{1}{6\sqrt{n}} \sum_{j=0}^3 w_j \Pr\{\chi_{r+2j}^2(\delta) < c_0^\phi(\alpha)\} \\ &- \frac{1}{72n} \sum_{j=0}^6 r_j \Pr\{\chi_{r+2j}^2(\delta) < c_0^\phi(\alpha)\}, \end{aligned}$$

where $c_0^\phi(\alpha)$ is the size α critical value for the ϕ -divergence statistic C_ϕ . Theorem 3.1 and the approximation (3.5) are generalizations of the results of Taneichi *et al.* which discussed the power divergence statistic.

4. Applications and numerical investigations

In this section, we consider $C_{\phi_a^R}$ given by (1.3) and $C_{\phi_a^L}$ given by (1.4) as concrete ϕ -divergence statistics. From here, we call them R_a statistic and L_a statistic, respectively, for short. We apply approximation (3.5) to the statistics. For each statistic, we investigate the accuracy of approximation (3.5), numerically. We refer to the approximation (3.5) as AE approximation. If we put $T^3 = 3(a-1)$ and $T^4 = 12(a-1)^2$ in Theorem 3.1, we can obtain the coefficients $w_i (i = 0, \dots, 3)$ and $r_j (j = 0, \dots, 6)$ of the AE approximation for R_a statistic. Similarly, if we put $T^3 = -(a+1)$ and $T^4 = 2(a^2 + a + 1)$ in Theorem 3.1, we can obtain the coefficients of the AE approximation for L_a statistic.

In order to investigate the performance of the AE approximation, we compare the approximation with the true power of the randomized size- α test that is calculated by direct enumeration (e.g. West and Kempthorne (1971) and Read and Cressie (1988, pp. 76–77)). The procedure for the comparison is as follows:

First, we fix the values of k , n , δ , and significance level. Second, we select null hypothesis and alternative hypotheses. Setting $p_j = 1/k$ for $j = 1, \dots, k$ we consider null hypothesis (1.2). We randomly generate vectors $\mathbf{c}^{(\alpha)} = (c_1^{(\alpha)}, \dots, c_k^{(\alpha)})'$, ($\alpha = 1, 2, \dots$) such that $\sum_{j=1}^k c_j^{(\alpha)} = 0$ and $\sum_{j=1}^k (c_j^{(\alpha)})^2 / p_j = \delta$. Excluding $\mathbf{c}^{(\alpha)}$ which does not satisfy $c_j^{(\alpha)} > -\sqrt{n}p_j$, ($j = 1, \dots, k$), we obtain new 500 vectors $\mathbf{c}^{(i)} = (c_1^{(i)}, \dots, c_k^{(i)})'$, ($i = 1, \dots, 500$). After that, we consider the alternative $H_{1,n}^{(i)} : \pi_j = p_j + c_j^{(i)} / \sqrt{n}$, ($j = 1, \dots, k$) corresponding to $\mathbf{c}^{(i)}$ for each i . Third, under each alternative $H_{1,n}^{(i)}$, we calculate the true power $PE(i)$, ($i = 1, \dots, 500$) of the randomized size- α test and the approximated power $P(i)$,

Table 1. M_1 and M_2 for the AE approximation ($k = 4$) of R_a statistic.

$a \backslash n$	M_1						M_2					
	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$
0.0	0	322	59	341	288	500	0	451	121	500	397	500
0.1	73	306	280	426	192	500	193	500	500	500	353	500
0.2	500	312	500	500	335	500	500	426	500	500	429	500
0.3	500	297	500	500	422	500	500	400	500	500	500	500
0.4	0	206	500	500	500	500	500	359	500	500	500	500
0.5	319	233	500	412	500	500	500	447	500	500	500	500
0.6	500	398	500	422	500	500	500	500	500	500	500	500
0.7	500	500	500	500	500	500	500	500	500	500	500	500
0.8	396	341	500	483	483	500	489	500	500	500	500	500
0.9	500	500	500	500	500	500	500	500	500	500	500	500
1.0	0	269	55	500	500	464	322	500	476	500	500	500

Table 2. M_1 and M_2 for the AE approximation ($k = 5$) of R_a statistic.

$a \backslash n$	M_1						M_2					
	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$
0.0	17	0	139	287	288	438	29	0	375	421	463	500
0.1	187	0	463	322	489	484	421	10	500	433	500	500
0.2	317	106	465	301	500	484	466	254	500	407	500	500
0.3	386	288	456	404	500	426	490	401	500	474	500	484
0.4	422	420	447	483	500	490	459	475	490	500	500	500
0.5	413	462	445	500	478	500	451	500	481	500	500	500
0.6	388	466	494	500	497	500	455	500	500	500	500	500
0.7	483	500	500	499	500	500	500	500	500	500	500	500
0.8	500	500	500	500	500	500	500	500	500	500	500	500
0.9	371	500	500	500	500	500	457	500	500	500	500	500
1.0	35	394	458	346	500	469	341	470	500	500	500	500

($i = 1, \dots, 500$) by the AE approximation. The corresponding critical value of the true power is also used for the AE approximation. We investigate the performance of the approximation based on the following two indices.

M_1 : The number of alternatives which satisfy $|P(i) - PE(i)| < 0.02$.

M_2 : The number of alternatives which satisfy $|P(i) - PE(i)| < 0.04$.

For the number of cells $k = 4, 5$, sample size $n = 5k, 6k, 7k, 8k, 9k, 10k$, noncentrality parameter $\delta = k$, and significance level $\alpha = 0.05$, we consider statistics R_a and L_a with $a = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$. The values of the indices M_1 and M_2 for $k = 4$ and $k = 5$ of the R_a statistic are listed in Table 1 and Table 2, respectively. The values of the indices M_1 and M_2 for $k = 4$ and $k = 5$ of the L_a statistic are listed in Table 3 and Table 4, respectively. In the tables, we use different sets of 500 alternatives for each numerical value.

From Tables 1 and 2, we find that the AE approximation satisfies $M_1 > 280$ and $M_2 > 390$ for almost all values of a when $n \geq 8k$ (except for $a = 0.1, n = 9k$).

Table 3. M_1 and M_2 for the AE approximation ($k = 4$) of L_a statistic.

$a \backslash n$	M_1						M_2					
	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$
0.0	500	159	500	499	500	456	500	404	500	500	500	500
0.1	500	164	500	368	500	500	500	426	500	492	500	500
0.2	468	178	500	231	500	500	500	409	500	452	500	500
0.3	424	117	429	290	500	500	500	367	500	431	500	500
0.4	500	88	500	373	500	500	500	293	500	481	500	500
0.5	410	125	500	407	500	500	500	297	500	500	500	500
0.6	475	86	500	483	500	500	500	237	500	500	500	500
0.7	493	158	500	500	496	500	500	325	500	500	500	500
0.8	494	275	500	500	500	500	500	381	500	500	500	500
0.9	234	145	417	500	397	500	500	318	500	500	500	500
1.0	81	148	312	500	319	500	498	325	500	500	456	500

Table 4. M_1 and M_2 for the AE approximation ($k = 5$) of L_a statistic.

$a \backslash n$	M_1						M_2					
	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$	$5k$	$6k$	$7k$	$8k$	$9k$	$10k$
0.0	332	413	434	490	462	500	405	488	472	500	500	500
0.1	283	317	416	488	442	500	401	464	462	500	495	500
0.2	294	368	412	487	443	495	408	486	459	500	495	500
0.3	311	285	401	472	453	500	417	463	455	500	500	500
0.4	283	400	423	436	457	479	413	491	463	500	500	500
0.5	291	381	425	433	449	451	419	457	464	500	500	500
0.6	259	377	406	441	466	446	419	448	464	500	500	500
0.7	225	343	398	330	488	442	415	437	466	473	500	500
0.8	366	263	395	372	478	396	446	386	465	465	500	480
0.9	368	91	435	332	495	397	446	272	483	437	500	471
1.0	292	14	429	312	493	397	435	162	495	417	500	465

From Tables 3 and 4, we also find that the AE approximation satisfies $M_1 > 310$ and $M_2 > 410$ for almost all value a when $n \geq 7k$ (except for $a = 0.2$, $n = 8k$ and $a = 0.3$, $n = 8k$). On the whole, the AE approximation performs well on the L_a and R_a statistics except for special cases.

We next compare the performance of the AE approximation with that of the other power approximations. Drost *et al.* (1989) proposed two power approximations A^a and B^a for the power divergence statistic $C_{\phi_a^P}$. The A^a approximation is constructed from a linear combination of mutually independent noncentral chi-square random variables. The B^a approximation is constructed from a linear expression of a single noncentral chi-square random variable which has the same moment as A^a . Sekiya *et al.* (1999) proposed NT approximation, which is a normal approximation based on normalizing transformation of the power divergence statistic $C_{\phi_a^P}$. On the basis of A^a and B^a approximations, we construct approximations for ϕ -divergence family of statistics C_ϕ and we call them A^ϕ and

B^ϕ approximations. Similarly, on the basis of the method of NT approximation for the power divergence statistic, we construct approximations for ϕ -divergence family of statistics C_ϕ and we call it N^ϕ approximation.

We describe the expressions of A^ϕ , B^ϕ , and N^ϕ approximations to the power of C_ϕ against alternative

$$H_1 : \boldsymbol{\pi} = \boldsymbol{\tau}, \quad (\boldsymbol{\tau} \neq \boldsymbol{p}),$$

where $\boldsymbol{\tau} = (\tau_1, \dots, \tau_k)'$ and $\boldsymbol{p} = (p_1, \dots, p_k)'$ is given by (1.2). Let

$$\begin{aligned} A_j &= \phi' \left(\frac{\tau_j}{p_j} \right), \quad (j = 1, \dots, k) \\ B_j &= \frac{1}{p_j} \phi'' \left(\frac{\tau_j}{p_j} \right), \quad (j = 1, \dots, k) \\ \boldsymbol{\tau}^{1/2} &= (\tau_1^{1/2}, \dots, \tau_k^{1/2})', \\ \boldsymbol{Q} &= \text{diag}((B_1 \tau_1)^{1/2}, \dots, (B_k \tau_k)^{1/2}), \end{aligned}$$

and

$$\boldsymbol{\mu} = n^{1/2} \left(\frac{A_1}{B_1 \sqrt{\tau_1}}, \dots, \frac{A_k}{B_k \sqrt{\tau_k}} \right)'.$$

Let $\gamma_1, \dots, \gamma_r$ be the non-zero eigenvalues and $\boldsymbol{S}$ the $k \times k$ orthogonal matrix of eigenvectors of $\boldsymbol{Q}(\boldsymbol{E} - \boldsymbol{\tau}^{1/2}(\boldsymbol{\tau}^{1/2})')\boldsymbol{Q}$:

$$\boldsymbol{S}'\boldsymbol{Q}(\boldsymbol{E} - \boldsymbol{\tau}^{1/2}(\boldsymbol{\tau}^{1/2})')\boldsymbol{Q}\boldsymbol{S} = \text{diag}(\gamma_1, \dots, \gamma_r, 0),$$

where $\boldsymbol{E}$ denotes the k -dimensional unit matrix. And let

$$\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)' = \boldsymbol{S}'\boldsymbol{Q}\boldsymbol{\mu}.$$

Then the A^ϕ approximation is represented as

$$(4.1) \quad 1 - \Pr \left\{ \sum_{j=1}^{k-1} \gamma_j \chi_1^2 \left(\frac{(\xi_j)^2}{\gamma_j} \right) < c_0^\phi(\alpha) - 2nD_\phi(\boldsymbol{\tau}, \boldsymbol{p}) + \sum_{j=1}^{k-1} (\xi_j)^2 \right\},$$

where $D_\phi(\cdot, \cdot)$ is given by (1.1). Let

$$\bar{B} = \frac{1}{k-1} \sum_{j=1}^k B_j \tau_j (1 - \tau_j)$$

and

$$\omega = n\bar{B}^{-2} \left\{ \sum_{j=1}^k A_j^2 \tau_j - \left(\sum_{j=1}^k A_j \tau_j \right)^2 \right\}.$$

Then the B^ϕ approximation is represented as

$$(4.2) \quad 1 - \Pr \left\{ \chi_r^2(\omega) < \frac{c_0^\phi(\alpha) - 2nD_\phi(\boldsymbol{\tau}, \boldsymbol{p}) + \bar{B}\omega}{\bar{B}} \right\}.$$

Table 5. M_1 for each approximation ($n = 8k$) of R_a statistic.

a	$k = 4$					$k = 5$				
	χ^2	A^ϕ	B^ϕ	N^ϕ	AE	χ^2	A^ϕ	B^ϕ	N^ϕ	AE
0.0	0	157	118	80	341	0	44	66	225	287
0.1	0	215	191	32	426	0	0	105	358	322
0.2	16	334	253	23	500	1	0	59	423	301
0.3	75	470	419	101	500	32	0	145	246	404
0.4	122	419	500	123	500	65	35	306	226	483
0.5	138	196	246	443	412	153	260	490	128	500
0.6	219	312	304	367	422	243	500	462	0	500
0.7	296	472	396	265	500	297	396	333	57	499
0.8	380	500	500	11	483	340	445	381	0	500
0.9	448	500	402	124	500	302	500	444	0	500
1.0	500	500	456	52	500	187	500	400	0	346

Table 6. M_2 for each approximation ($n = 8k$) of R_a statistic.

a	$k = 4$					$k = 5$				
	χ^2	A^ϕ	B^ϕ	N^ϕ	AE	χ^2	A^ϕ	B^ϕ	N^ϕ	AE
0.0	0	371	236	500	500	0	109	177	500	421
0.1	8	467	323	500	500	0	0	287	483	433
0.2	60	500	416	500	500	15	0	264	465	407
0.3	131	500	500	500	500	66	382	500	500	474
0.4	285	500	500	500	500	169	500	500	500	500
0.5	301	420	466	500	500	291	500	500	500	500
0.6	350	435	426	500	500	387	500	500	499	500
0.7	465	500	500	500	500	423	500	466	500	500
0.8	500	500	500	476	500	456	500	468	500	500
0.9	500	500	500	500	500	489	500	500	350	500
1.0	500	500	500	500	500	500	500	500	43	500

Let

$$M_{jl}^{(2)} = \delta_{jl}\tau_j - \tau_j\tau_l,$$

$$M_{jlm}^{(3)} = \delta_{jl}\delta_{jm}\tau_j - (\delta_{jl}\tau_j\tau_m + \delta_{jm}\tau_j\tau_l + \delta_{lm}\tau_j\tau_l) + 2\tau_j\tau_l\tau_m,$$

$$\nu = D_\phi(\tau, \mathbf{p}),$$

$$\nu_1 = \frac{1}{2} \sum_{j=1}^k B_j M_{jj}^{(2)},$$

$$\nu_2 = \sum_{j=1}^k \sum_{l=1}^k A_j A_l M_{jl}^{(2)},$$

$$\nu_3 = \sum_{j=1}^k \sum_{l=1}^k \sum_{m=1}^k (A_j A_l A_m M_{jlm}^{(3)} + 3A_j A_l B_m M_{jm}^{(2)} M_{lm}^{(2)}),$$

Table 7. M_1 for each approximation ($n = 8k$) of L_a statistic.

a	$k = 4$					$k = 5$				
	χ^2	A^ϕ	B^ϕ	N^ϕ	AE	χ^2	A^ϕ	B^ϕ	N^ϕ	AE
0.0	238	496	496	0	499	199	392	392	1	490
0.1	154	184	187	449	368	185	385	450	12	488
0.2	109	16	0	484	231	172	400	499	1	487
0.3	107	91	85	449	290	139	260	454	60	472
0.4	120	169	224	452	373	87	28	76	319	436
0.5	119	255	391	404	407	86	115	305	212	433
0.6	120	383	495	209	483	59	56	286	179	441
0.7	125	485	500	38	500	36	0	78	416	330
0.8	111	500	469	9	500	33	0	142	260	372
0.9	73	470	409	92	500	27	0	127	271	332
1.0	44	367	308	72	500	14	0	134	320	312

Table 8. M_2 for each approximation ($n = 8k$) of L_a statistic.

a	$k = 4$					$k = 5$				
	χ^2	A^ϕ	B^ϕ	N^ϕ	AE	χ^2	A^ϕ	B^ϕ	N^ϕ	AE
0.0	396	500	500	500	500	336	500	500	500	500
0.1	307	437	449	500	492	319	500	500	500	500
0.2	234	380	401	500	452	309	500	500	498	500
0.3	242	367	398	500	431	277	500	500	500	500
0.4	258	417	493	500	481	219	500	500	500	500
0.5	254	457	500	500	500	198	500	500	500	500
0.6	259	500	500	500	500	169	500	500	500	500
0.7	252	500	500	500	500	96	460	500	500	473
0.8	199	500	500	500	500	86	455	500	500	465
0.9	122	500	500	500	500	55	370	500	500	437
1.0	94	500	500	500	500	32	302	465	500	417

$$\zeta = -\frac{\nu\nu_3}{3(\nu_2)^2},$$

and

$$\psi(x) = \begin{cases} \frac{\nu}{\zeta + 1} \left(\left(\frac{x}{\nu} \right)^{\zeta+1} - 1 \right) & (\zeta \neq -1) \\ \nu \ln \frac{x}{\nu} & (\zeta = -1). \end{cases}$$

Then the N^ϕ approximation is represented as

$$(4.3) \quad 1 - \Phi \left(\sqrt{\frac{n}{\nu_2}} \left\{ \psi \left(\frac{c_0^\phi(\alpha)}{2n} \right) - \frac{1}{n} \left(\nu_1 - \frac{\nu_3}{6\nu_2} \right) \right\} \right),$$

where $\Phi(\cdot)$ denotes the standard normal distribution function. If we consider ϕ_a^R as ϕ in (4.1), (4.2), and (4.3), we obtain A^ϕ , B^ϕ , and N^ϕ approximations for R_a statistic, respectively. Similarly, if we consider ϕ_a^L as ϕ in (4.1), (4.2), and (4.3), we obtain A^ϕ , B^ϕ , and N^ϕ approximations for L_a statistic, respectively. When

we put

$$\tau_j = p_j + \frac{c_j}{\sqrt{n}}, \quad (j = 1, \dots, k),$$

we regard the A^ϕ , the B^ϕ , and the N^ϕ approximations as approximations to the power of statistics against $H_{1,n}$.

In order to compare the performance of the AE approximation with the A^ϕ approximation, the B^ϕ approximation, the N^ϕ approximation, and the ordinary χ^2 approximation, i.e., $1 - \Pr\{\chi_r^2(\delta) < c_0^\phi(\alpha)\}$, we investigate the performance of the other approximations in a similar fashion to that of the AE approximation. Tables 5 and 6 illustrate the values of the indices M_1 and M_2 for each approximation of the R_a statistic when $n = 8k$, respectively. Tables 7 and 8 illustrate the values of the indices M_1 and M_2 for each approximation of the L_a statistic. In tables, we use different sets of 500 alternatives for each statistic.

From Tables 5–8, we find that the AE approximation satisfies $M_1 > 230$ and $M_2 > 340$ for every statistics. On the other hand, all of the other approximations indicate very bad values ($M_1 = 0$ or $M_2 < 180$) for some statistics. Furthermore, from more extensive investigations of the accuracy of the approximations, we find that the performance of the AE approximation is very stable for every statistics, while the performance of the other approximations depend on the statistics used. Then we can conclude that the AE approximation totally performs better than the other approximations when $n \geq 8k$ both on the L_a and the R_a statistics.

5. Concluding remarks

In this paper, omitting the discontinuous term, we proposed an asymptotic approximation for the distribution of C_ϕ under $H_{1,n}$. This approximation is a generalization of the approximation for the power divergence family of statistics given by Taneichi *et al.*. For the power divergence family of statistics, they investigated the performance of the proposed approximation, and compared it with the other approximations, numerically. They concluded that the error of the proposed approximation is not large on the whole. They also concluded that the proposed approximation performs better than the other approximations.

By the generalization, we can also derive the asymptotic approximations for the R_a and the L_a statistics. From the numerical investigations given in Section 4, for the R_a and the L_a statistics, we arrived at the similar conclusion of the power divergence family of statistics.

Acknowledgements

The authors are very grateful to the referees for their valuable comments.

REFERENCES

- Ali, S. M. and Silvey, S. D. (1966). A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. B*, **28**, 131–142.
- Bhattacharya, R. N. and Ranga Rao, R. (1976). *Normal Approximation and Asymptotic Expansions*, Wiley.

- Burbea, J. and Rao, C. R. (1982). On the convexity of some divergence measures based on entropy functions, *IEEE Trans. Information Theory*, **28**, 489–495.
- Cressie, N. and Read, T. R. C. (1984). Multinomial goodness-of-fit tests, *J. R. Statist. Soc. B*, **46**, 440–464.
- Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations, *Studia Sci. Math. Hungar.*, **2**, 299–318.
- Drost, F. C., Kallenberg, W. C. M., Moore, D. S. and Oosterhoff, J. (1989). Power approximations to multinomial tests of fit, *J. Amer. Statist. Assoc.*, **84**, 130–141.
- Esséen, C. G. (1945). Fourier analysis of distribution functions: A mathematical study of the Laplace-Gaussian law, *Acta Math.*, **77**, 1–125.
- Lin, J. (1991). Divergence measures based on the Shannon entropy, *IEEE Trans. Information Theory*, **37**, 145–151.
- Menéndez, M. L., Pardo, J. A., Pardo, L. and Pardo, M. C. (1997). Asymptotic approximations for the distributions of the (h, ϕ) -divergence goodness-of-fit statistics: application to Renyi's statistic, *Kybernetes*, **26**(4), 442–452.
- Pardo, L., Pardo, M. C. and Zografos, K. (1999). Homogeneity for multinomial populations based on ϕ -divergences, *J. Japan Statist. Soc.*, **29**, 213–228.
- Ranga Rao, R. (1961). On the central limit theorem in R_k , *Bull. Amer. Math. Soc.*, **67**, 359–361.
- Read, T. R. C. (1984). Closer asymptotic approximations for the distributions of the power divergence goodness-of-fit statistics, *Ann. Inst. Statist. Math.*, **36**, 59–69.
- Read, T. R. C. and Cressie, N. A. C. (1988). *Goodness-of-fit statistics for discrete multivariate data*, Springer-Verlag.
- Rukhin, A. L. (1994). Optimal estimator for the mixture parameter by the method of moments and information affinity, in *Trans. 12th Prague Conference on Information Theory*, 214–219.
- Sekiya, Y. (1990). Some properties of goodness-of-fit statistics based on the J-divergence (in Japanese), *KUSHIRO RONSHU*, No. **22**, 29–43.
- Sekiya, Y., Taneichi, N. and Imai, H. (1999). A power approximation for the multinomial goodness-of-fit test based on a normalizing transformation, *J. Japan Statist. Soc.*, **29**, 79–87.
- Siotani, M. and Fujikoshi, Y. (1984). Asymptotic approximations for the distributions of multinomial goodness-of-fit statistics, *Hiroshima Math. J.*, **14**, 115–124.
- Taneichi, N. (1993). Goodness-of-fit test based on the J-divergence for the multinomial distribution, *J. Japan Statist. Soc.*, **23**, 9–18.
- Taneichi, N. (1995). *Theory of discrete multivariate analysis and its applications* (in Japanese), Doctorial dissertation, Division of Information Engineering, Hokkaido University.
- Taneichi, N. and Sato, Y. (1989). Goodness-of-fit test for multinomial distribution by J-divergence, *Preprints of the "Symposium on the Analysis of Statistical Information"*, 49–63.
- Taneichi, N., Sekiya, Y. and Suzukawa, A. Asymptotic approximations for the distributions of the multinomial goodness-of-fit statistics under local alternatives, *Journal of Multivariate Analysis* (to appear).
- Tomizawa, S., Seo, T. and Yamamoto, H. (1998). Power-divergence type measure of departure from symmetry for square contingency tables that have nominal categories, *J. Applied Statistics*, **25**, 387–398.
- West, E. N. and Kempthorne, O. (1971). A comparison of the χ^2 and likelihood ratio tests for composite alternatives, *J. Statist. Comput. Simul.*, **1**, 1–33.
- Yarnold, J. K. (1972). Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set, *Ann. Math. Statist.*, **43**, 1566–1580.
- Zografos, K., Ferentions, K. and Papaioannou, T. (1990). ϕ -divergence statistics: Sampling properties and multinomial goodness-of-fit and divergence tests, *Commun. Statist.-Theory Meth.*, **19**, 1785–1802.