Ann. Inst. Statist. Math.
Vol. 55, No. 4, 831–848 (2003)

IMPROVEMENTS OF GOODNESS-OF-FIT STATISTICS FOR SPARSE MULTINOMIALS BASED ON NORMALIZING TRANSFORMATIONS

NOBUHIRO TANEICHI[1], YURI SEKIYA[2] AND HIDEYUKI IMAI[3]

[1] *Obihiro University of Agriculture and Veterinary Medicine, Inada-cho, Obihiro 080-8555, Japan*
[2] *Hokkaido University of Education, Kushiro Campus, Shiroyama, Kushiro 085-8580, Japan*
[3] *Division of Systems and Information Engineering, Hokkaido University, Sapporo 060-8628, Japan*

(Received March 4, 2002; revised November 15, 2002)

Abstract. We consider multinomial goodness-of-fit tests for a specified simple hypothesis under the assumption of sparseness. It is shown that the asymptotic normality of the Pearson X^2 statistic (X_k^2) and the log-likelihood ratio statistic (G_k^2) assuming sparseness. In this paper, we improve the asymptotic normality of X_k^2 and G_k^2 statistics based on two kinds of normalizing transformation. The performance of the transformed statistics is numerically investigated.

Key words and phrases: Normalizing transformation, Pearson X^2 statistic, log-likelihood ratio statistic, sparse multinomials.

1. Introduction

When we consider goodness-of-fit tests for a multinomial distribution, we usually use the Pearson X^2 statistic (X_k^2) or the log-likelihood ratio statistic (G_k^2). When the null hypothesis is true, both statistics have the same limiting central chi-squared distribution under the traditional limiting argument that requires that the number of cells k is fixed and that all of the expected cell frequencies tend to infinity as the sample size n tends to infinity. However, the condition is not always satisfied in practice (Koehler and Larntz (1980)). For such cases, another asymptotic theory is needed. For testing a specified simple hypothesis, Morris (1975) and Holst (1972) have shown the asymptotic normality of X_k^2 and G_k^2 under different conditions, respectively, when both n and k tend to infinity while all of the expected cell frequencies remain finite. We refer to these as sparseness assumptions. Koehler and Larntz (1980) numerically showed that the limiting normal approximation is generally more accurate for G_k^2 than for X_k^2.

In order to improve the approximation to the distributions of the goodness-of-fit test statistics for a multinomial distribution, Yarnold (1972), Siotani and Fujikoshi (1984), and Read (1984) studied asymptotic expansion under the null hypothesis for fixed k. Furthermore, Taneichi *et al.* (2002*a*) studied asymptotic expansion under local alternatives. However, even for an ordinary multinomial case (not assuming sparseness), complete asymptotic expansions for the distributions of the statistics have not yet been derived, since it is very difficult to evaluate the order of discontinuous terms. Under the sparseness assumptions, even a formal asymptotic expansion has not yet been derived. In order to construct a normalizing transformation based on a concept similar to that of the Bartlett transformation (Hall (1992) and Yanagihara and Tonda (2003)), Edgeworth expansion of the distribution function of the statistics up to the order of $k^{-1/2}$ is needed. Thus, the Bartlett type of transformation can not be used to improve X_k^2 and G_k^2 statistics under the sparseness assumptions. In this paper, we show how the asymptotic normality of X_k^2 and G_k^2 statistics can be improved by using other kinds of normalizing transformation. In Section 2, we introduce the theory of a multinomial goodness-of-fit test for a specified simple hypothesis under the sparseness assumptions and the theory of the normalizing transformation that we use for improvement. In Section 3, we show how normalizing transformations of X_k^2 are derived, and in Section 4, we show how normalizing transformations of G_k^2 are derived. The performance of the improved statistics is examined in Section 5.

2. Preliminaries

We first introduce the theory of a multinomial goodness-of-fit test for a specified simple null hypothesis under the sparseness assumptions. Morris (1975) derived a certain limit theorem for sums of functions of multinomial frequencies. By applying this theorem, he showed asymptotic normality of the Pearson X^2 statistic and the log-likelihood ratio statistic under the simple null hypothesis as follows.

Consider a sequence of multinomial distributions with increasing numbers of categories. For the k-th multinomial in the sequence, let $\boldsymbol{X}_k = (X_{1k}, X_{2k}, \ldots, X_{kk})'$ be distributed as $\mathrm{Mult}_k(n_k, \boldsymbol{p}_k)$, where $\boldsymbol{p}_k = (p_{1k}, p_{2k}, \ldots, p_{kk})'$. We consider the simple null hypothesis

$$H_0 : \boldsymbol{p}_k = \boldsymbol{p}_{0k},$$

where $\boldsymbol{p}_{0k} = (p_{01k}, p_{02k}, \ldots, p_{0kk})'$. We assume that

$$n_k \to \infty \quad \text{as} \quad k \to \infty, \tag{2.1}$$

and we place the following restrictions on $\boldsymbol{p}_k$ and $\boldsymbol{p}_{0k}$:

$$p_{0jk} > 0 \quad \text{for all} \quad j, k,$$
$$\sum_{j=1}^{k} p_{0jk} = 1,$$
$$\max_{1 \le j \le k} p_{jk} = o(1) \quad \text{as} \quad k \to \infty,$$

and

$$n_k p_{jk} \ge \varepsilon \quad \text{for all} \quad j, k, \quad \text{some} \quad \varepsilon > 0 \quad \text{fixed.} \tag{2.2}$$

Consider the limiting distribution of the Pearson X^2 statistic

$$X_k^2 = \sum_{j=1}^{k} \frac{(X_{jk} - n_k p_{0jk})^2}{n_k p_{0jk}}$$

and the log-likelihood ratio statistic

$$G_k^2 = 2 \sum_{j=1}^{k} X_{jk} \log\left(\frac{X_{jk}}{n_k p_{0jk}}\right),$$

respectively, when $n_k \to \infty$ as $k \to \infty$. Let

$$\mu_{P,k} = k$$

and

$$s_{P,k}^2 = 2k + \frac{1}{n_k}\left(\sum_{j=1}^{k}\frac{1}{p_{0jk}} - k^2\right).$$

Then under H_0,

$$\frac{X_k^2 - \mu_{P,k}}{s_{P,k}} \xrightarrow{L} N(0,1) \quad \text{as} \quad k \to \infty, \tag{2.3}$$

where $\xrightarrow{L}$ denotes convergence in the distribution. Let

$$\mu_{LR,k} = 2\sum_{j=1}^{k} E[I(Y_{jk}, n_k p_{0jk})] \tag{2.4}$$

and

$$s_{LR,k}^2 = 4\sum_{j=1}^{k} V[I(Y_{jk}, n_k p_{0jk})] - n_k\gamma_k^2, \tag{2.5}$$

where $Y_{jk}(j = 1, \ldots, k)$ are mutually independent Poisson random variables such that $E(Y_{jk}) = E(X_{jk})$.

$$I(y,m) = \begin{cases} y\log\left(\dfrac{y}{m}\right) - y + m, & \text{if} \quad y > 0 \\ m, & \text{if} \quad y = 0 \end{cases}$$

is the Kullback-Leibler information kernel for a Poisson random variable y with mean m, and

$$\gamma_k = \frac{2}{n_k}\sum_{j=1}^{k}\mathrm{Cov}[I(Y_{jk}, n_k p_{0jk}), Y_{jk}]. \tag{2.6}$$

Then under H_0,

$$\frac{G_k^2 - \mu_{LR,k}}{s_{LR,k}} \xrightarrow{L} N(0,1) \quad \text{as} \quad k \to \infty. \tag{2.7}$$

Next, we introduce a formula for normalizing transformations derived by Taneichi *et al.* (2002*b*). This formula is based on the general framework for constructing transformation given by Konishi (1981, 1991)

Formula 1. Let T_n be a random variable whose distribution depends on the parameter n. We assume that the mean, variance, and third moment about the mean can be expanded as

$$E(T_n) = \mu + \frac{1}{n}\mu_1 + o\left(\frac{1}{n}\right),$$

$$V(T_n) = \frac{1}{n}\sigma^2 + o\left(\frac{1}{n}\right),$$

and

$$E[\{T_n - E(T_n)\}^3] = \frac{1}{n^2}\nu + o\left(\frac{1}{n^2}\right),$$

respectively. We also assume that

$$\frac{\sqrt{n}(T_n - \mu)}{\sigma} \xrightarrow{L} N(0,1), \qquad \text{as} \quad n \to \infty.$$

Let

$$g_1(T_n) = \begin{cases} \dfrac{\sqrt{n}}{\sigma}\left[\dfrac{\mu}{\eta}\left\{\left(\dfrac{T_n}{\mu}\right)^\eta - 1\right\} - \dfrac{1}{n}\left(\dfrac{1}{2}\sigma^2\xi + \mu_1\right)\right] & (\eta \neq 0) \\ \dfrac{\sqrt{n}}{\sigma}\left[\mu \log\dfrac{T_n}{\mu} - \dfrac{1}{n}\left(\dfrac{1}{2}\sigma^2\xi + \mu_1\right)\right] & (\eta = 0) \end{cases}$$

and

$$g_2(T_n) = \frac{\sqrt{n}}{\sigma}\left\{\frac{1}{\xi}\left(e^{\xi(T_n - \mu)} - 1\right) - \frac{1}{n}\left(\frac{1}{2}\sigma^2\xi + \mu_1\right)\right\},$$

where

$$\xi = -\frac{\nu}{3\sigma^4}$$

and

$$\eta = \xi\mu + 1.$$

If T_n is continuous, then $g_i (i = 1, 2)$ are normalizing transformations in the sense that

$$P(g_i(T_n) < x) = \Phi(x) + o\left(\frac{1}{\sqrt{n}}\right), \qquad (i = 1, 2),$$

where Φ is the standard normal distribution function. If T_n is discrete, then $g_i (i = 1, 2)$ are normalizing transformations in the sense that

$$g_i(T_n) \xrightarrow{L} N(0,1), \qquad \text{as} \quad n \to \infty, \qquad (i = 1, 2)$$

and

$$E[\{g_i(T_n) - E(g_i(T_n))\}^3] = o\left(\frac{1}{\sqrt{n}}\right), \qquad (i = 1, 2).$$

3. Normalizing transformations of the X_k^2 goodness-of-fit statistic

In this section, we show how normalizing transformations of the Pearson X^2 goodness-of-fit statistic for sparse multinomials are derived. From here on, we further assume that

(3.1) $$n_k \to \infty \quad \text{as} \quad k \to \infty \quad \text{so that} \quad \lim_{k\to\infty} \frac{n_k}{k} = c \quad (0 < c < \infty).$$

Assumption (3.1) was used by Holst (1972). Put

$$T_k = \frac{X_k^2}{k},$$

then the mean, variance, and third moment about the mean of T_k under H_0 are

$$E(T_k) = 1 - \frac{1}{k}, \tag{3.2}$$

$$V(T_k) = \frac{1}{k}\sigma_T^2 + o\left(\frac{1}{k}\right), \quad \text{as} \quad k \to \infty, \tag{3.3}$$

and

$$E[\{T_k - E(T_k)\}^3] = \frac{1}{k^2}\nu_T + o\left(\frac{1}{k^2}\right), \quad \text{as} \quad k \to \infty, \tag{3.4}$$

where

$$\sigma_T^2 = 2 + d - \frac{1}{c},$$

$$\nu_T = 8 + 22d + e - \frac{3(6+d)}{c} + \frac{2}{c^2},$$

$$d = \lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \frac{1}{n_k p_{0jk}}, \tag{3.5}$$

$$e = \lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} \frac{1}{(n_k p_{0jk})^2}, \tag{3.6}$$

and c is given by (3.1). From assumption (2.2), it is clear that $e < \infty$ and $d < \infty$. Therefore, the values of σ_T^2 and ν_T are finite. Since

$$\frac{\sqrt{k}(T_k - 1)}{\sigma_T} = \left(\frac{X_k^2 - \mu_{P,k}}{s_{P,k}}\right)\left(\frac{s_{P,k}}{\sqrt{k}\sigma_T}\right)$$

and

$$\lim_{k\to\infty} \frac{s_{P,k}}{\sqrt{k}} = \sigma_T,$$

we obtain

$$\frac{\sqrt{k}(T_k - 1)}{\sigma_T} \xrightarrow{L} N(0,1), \quad \text{as} \quad k \to \infty \tag{3.7}$$

under H_0 from (2.3). Therefore, applying (3.2), (3.3), (3.4) and (3.7) to Formula 1, we find that

$$g_1(T_k) = \begin{cases} \dfrac{\sqrt{k}}{\sigma_T}\left\{\dfrac{1}{\eta_T}(T_k^{\eta_T} - 1) - \dfrac{1}{k}\left(\dfrac{1}{2}\sigma_T^2\xi_T - 1\right)\right\} & (\eta_T \neq 0) \\ \dfrac{\sqrt{k}}{\sigma_T}\left\{\log T_k - \dfrac{1}{k}\left(\dfrac{1}{2}\sigma_T^2\xi_T - 1\right)\right\} & (\eta_T = 0) \end{cases}$$

and

$$g_2(T_k) = \frac{\sqrt{k}}{\sigma_T}\left\{\frac{1}{\xi_T}\left(e^{\xi_T(T_k - 1)} - 1\right) - \frac{1}{k}\left(\frac{1}{2}\sigma_T^2\xi_T - 1\right)\right\}$$

are the normalizing transformations of T_k under H_0, where

$$\xi_T = -\frac{\nu_T}{3\sigma_T^4}$$

and

$$\eta_T = \xi_T + 1.$$

In the above discussion, c, d, and e are defined by the limit values. We must therefore estimate the values in practical applications. We propose that c, d, and e be estimated as $c^* = n_k/k$, $d^* = k^{-1}\sum_{j=1}^{k}(n_k p_{0jk})^{-1}$, and $e^* = k^{-1}\sum_{j=1}^{k}(n_k p_{0jk})^{-2}$, respectively.

Haldane (1937) derived exact moments of X_k^2 under the null hypothesis. Modified transformations of g_1 and g_2 can be considered by the exact moments. Using the exact moments, we can formally write

$$V(T_k) = \frac{1}{k}\sigma^2_{exact}$$

and

$$E[\{T_k - E(T_k)\}^3] = \frac{1}{k^2}\nu_{exact},$$

where

$$\sigma^2_{exact} = 2\left(1 - \frac{1}{k}\right) + \frac{1}{n_k k}\left\{R_1 - (k^2 + 2k - 2)\right\},$$

$$\nu_{exact} = 8\left(1 - \frac{1}{k}\right) + \frac{2}{n_k k}\left\{11R_1 - (9k^2 + 18k - 16)\right\}$$
$$+ \frac{1}{n_k^2 k}\left\{R_2 - (3k + 22)R_1 + 2(k^3 + 9k^2 + 14k - 12)\right\},$$

$R_1 = \sum_{j=1}^{k} p_{0jk}^{-1}$, and $R_2 = \sum_{j=1}^{k} p_{0jk}^{-2}$. We define the modified transformations g_{1E} and g_{2E} of T_k by substituting σ_{exact} and ν_{exact} for σ_T and ν_T in transformations g_1 and g_2, respectively.

4. Normalizing transformations of the G_k^2 goodness-of-fit statistic

In this section, we consider normalizing transformations of the G_k^2 goodness-of-fit statistic for sparse multinomials. Let

$$S_k = \frac{G_k^2}{k}.$$

Then by (2.4) and (2.5), the mean and variance of S_k under H_0 and assumption (3.1) are

$$E(S_k) = \mu_S + o(1), \quad \text{as} \quad k \to \infty \tag{4.1}$$

and

$$V(S_k) = \frac{1}{k}\sigma_S^2 + o\left(\frac{1}{k}\right), \quad \text{as} \quad k \to \infty, \tag{4.2}$$

where

$$\mu_S = \lim_{k\to\infty} \frac{\mu_{LR,k}}{k} \tag{4.3}$$

and

$$\sigma_S^2 = \lim_{k\to\infty} \frac{s^2_{LR,k}}{k}, \tag{4.4}$$

respectively. Next, we consider the third moment about the mean of S_k under H_0. When we represent the third moment about the mean of S_k as

$$E[\{S_k - E(S_k)\}^3] = \frac{1}{k^2}\nu_S + o\left(\frac{1}{k^2}\right), \quad \text{as} \quad k \to \infty, \tag{4.5}$$

we substitute

$$\nu_S' = \lim_{k\to\infty} \frac{\nu_{LR,k}}{k} \tag{4.6}$$

for ν_S, where

$$\nu_{LR,k} = 8k\mu_k^{3,0} - 12k\mu_k^{2,1}\gamma_k + 6k\mu_k^{1,2}\gamma_k^2 - n_k\gamma_k^3, \tag{4.7}$$

$$\mu_k^{l,m} = \frac{1}{k}\sum_{j=1}^{k} E\left[\{I(Y_{jk}, n_k p_{0jk}) - E[I(Y_{jk}, n_k p_{0jk})]\}^l (Y_{jk} - n_k p_{0jk})^m\right], \tag{4.8}$$

$Y_{jk} (j = 1, \ldots, k)$ are mutually independent Poisson random variables with $E(Y_{jk}) = n_k p_{0jk}$ and γ_k is given by (2.6). (The reason for substituting ν_S' for ν_S is given in Appendix A.) We can prove that the values of μ_S, σ_S^2, and ν_S' are all finite. (The proof is shown in Appendix B.) If we assume that $E(S_k)$ can be expanded as

$$E(S_k) = \mu_S + \frac{1}{k}\mu_{S1} + o\left(\frac{1}{k}\right), \quad \text{as} \quad k \to \infty \tag{4.9}$$

instead of (4.1), then

$$\sqrt{k}\left(\frac{\mu_{LR,k}}{k} - \mu_S\right) \to 0, \quad \text{as} \quad k \to \infty.$$

Therefore, since

$$\frac{\sqrt{k}(S_k - \mu_S)}{\sigma_S} = \left(\frac{G_k^2 - \mu_{LR,k}}{s_{LR,k}}\right)\left(\frac{s_{LR,k}}{\sqrt{k}\sigma_S}\right) + \frac{\sqrt{k}\left(\frac{\mu_{LR,k}}{k} - \mu_S\right)}{\sigma_S},$$

under H_0 and assumption (4.9), we find that

$$\frac{\sqrt{k}(S_k - \mu_S)}{\sigma_S} \xrightarrow{L} N(0,1), \quad \text{as} \quad k \to \infty, \tag{4.10}$$

by (2.7) and (4.4). Thus, applying (4.2), (4.6), assumption (4.9), and (4.10) to Formula 1, we find that

$$g_1(S_k) = \begin{cases} \dfrac{\sqrt{k}}{\sigma_S}\left[\dfrac{\mu_S}{\eta_S}\left\{\left(\dfrac{S_k}{\mu_S}\right)^{\eta_S} - 1\right\} - \dfrac{1}{k}\left(\dfrac{1}{2}\sigma_S^2\xi_S + \mu_{S1}\right)\right] & (\eta_S \neq 0) \\ \dfrac{\sqrt{k}}{\sigma_S}\left[\mu_S \log\dfrac{S_k}{\mu_S} - \dfrac{1}{k}\left(\dfrac{1}{2}\sigma_S^2\xi_S + \mu_{S1}\right)\right] & (\eta_S = 0) \end{cases}$$

and

$$g_2(S_k) = \frac{\sqrt{k}}{\sigma_S}\left\{\frac{1}{\xi_S}\left(e^{\xi_S(S_k - \mu_S)} - 1\right) - \frac{1}{k}\left(\frac{1}{2}\sigma_S^2\xi_S + \mu_{S1}\right)\right\}$$

are the normalizing transformations of S_k under H_0, where

$$\xi_S = -\frac{\nu'_S}{3\sigma_S^4}$$

and

$$\eta_S = \xi_S \mu_S + 1.$$

It is difficult to derive μ_{S1} defined by (4.9) analytically. Furthermore, μ_S, σ_S^2, and ν'_S are defined by the limit values. We must therefore estimate the values in practical applications. For μ_S, σ_S^2, and ν'_S, we propose the use of the estimates $\mu_S^* = \mu_{LR,k}/k$, $\sigma_S^{2*} = s_{LR,k}^2/k$ and $\nu_S^{'*} = \nu_{LR,k}/k$, respectively. For μ_{S1}, we propose the use of a simulated estimate defined by

$$\mu_{S1}^* = k(E^*(S_k) - \mu_S^*), \tag{4.11}$$

where $E^*(S_k)$ is a simulated value of $E(S_k)$ calculated by using multinomial random vectors.

5. Numerical investigation of the performance of the improved statistics

In this section, we numerically investigate the performance of the transformed statistics proposed in Sections 3 and 4. We consider the following three hypotheses for the investigation.

1. $\boldsymbol{p}_{0k} = (\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k})'$.
2. $\boldsymbol{p}_{0k} = (q_1, q_2, \ldots, q_k)'$, where $q_j = \frac{1}{k}(0.1 + 0.9c_j)$ and $c_j = \sum_{i=j}^{k} \frac{1}{i}$, $(j = 1, \ldots, k)$.
3. $\boldsymbol{p}_{0k} = (q_1, q_2, \ldots, q_k)'$, where $q_j = \frac{1.95}{k}$, $(j = 1, 2, \ldots, 0.5k)$, and $q_j = \frac{0.05}{k}$, $(j = 0.5k + 1, \ldots, k)$.

Hypothesis 1 is a symmetric hypothesis, and hypotheses 2 and 3 are typical cases considered by Koehler (1986). We also consider multinomial distributions with the number of cells $k = 4$, 10, 40, 100, and 400. For each cell size, sample sizes are selected such that $\lambda = n_k/k$ achieves the values 0.5, 1, 3, and 5. For each of the three null hypotheses selected and each combination of λ and k, we generate 100,000 multinomial random vectors and calculate the values of $(X_k^2 - \mu_{P,k})/s_{P,k}$, $g_1(T_k)$, $g_{1E}(T_k)$, $g_2(T_k)$, $g_{2E}(T_k)$, $(G_k^2 - \mu_{LR,k})/s_{LR,k}$, $g_1(S_k)$, and $g_2(S_k)$, respectively. For calculating the values of $g_1(T_k)$ and $g_2(T_k)$, we use the estimates c^*, d^*, and e^* for c, d, and e, respectively. Similarly, for calculating the values of $g_1(S_k)$ and $g_2(S_k)$, we use the estimates μ_S^*, σ_S^{2*}, and $\nu_S^{'*}$ for μ_S, σ_S^2, and ν'_S, respectively. The value of μ_{S1}^* defined by (4.11) is approximated by 10,000 multinomial random vectors. Let s_α be the simulated upper 100α-th percentile for the statistics, and let z_α be the upper 100α-th percentile of the standard normal distribution. We investigate the performance of the statistics based on the quantity $e_\alpha = s_\alpha - z_\alpha$. The values of $e_\alpha \times 10^5$ ($\alpha = 0.01$, 0.05, 0.1, 0.2) for each statistic and hypothesis are shown in Tables 1–6. The values of the index $Q = (n_k k)^{-1} \sum_{j=1}^{k} p_{0jk}^{-1}$ for each hypothesis and situation are also shown. In these tables, statistics $(X_k^2 - \mu_{P,k})/s_{P,k}$, $g_1(T_k)$, $g_{1E}(T_k)$, $g_2(T_k)$, $g_{2E}(T_k)$, $(G_k^2 - \mu_{LR,k})/s_{LR,k}$, $g_1(S_k)$, and $g_2(S_k)$ are denoted by the symbols X, g_1X, $g_{1E}X$, g_2X, $g_{2E}X$, G, g_1G, and g_2G, respectively.

In order to arrange the numerical results in Tables 1–6, we prepare other tables. Let $e_{\max}$ be the maximum value of $|e_\alpha|$ among $\alpha = 0.01$, 0.05, 0.1, and 0.2, where $|e_\alpha|$ denotes the absolute value of e_α. For each λ, statistic, and hypothesis, minimum values

Table 1. Values of $e_\alpha \times 10^5$ under hypothesis 1 for $\lambda = 5$ and $\lambda = 3$.

k	Q	α	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
			$\lambda = 5$							
4	0.20	0.2	−55878	3542	15109	3429	14950	−52554	3730	3585
		0.1	−57444	−6316	9744	−7676	7890	−51209	−3700	−5300
		0.05	−37206	−4199	15392	−10029	7721	−38459	−5789	−11242
		0.01	7781	−10526	11719	−33941	−17330	53754	14884	−17663
10	0.20	0.2	−39441	−3628	423	−3807	220	−31969	770	511
		0.1	−29768	−3764	2458	−5305	727	−20391	2125	279
		0.05	−12432	−1313	6469	−5997	1246	−5704	3140	−1822
		0.01	26749	−2718	7009	−19198	−11124	31915	2927	−13707
40	0.20	0.2	−17080	−1100	0	−1259	−164	−13864	397	228
		0.1	−7407	1596	3299	769	2448	−7290	832	55
		0.05	984	1760	3895	−186	1893	1416	1891	47
		0.01	22277	1105	3950	−4913	−2233	20336	1119	−4473
100	0.20	0.2	−7795	1614	2085	1516	1987	−7383	666	574
		0.1	−3704	1251	1958	856	1558	−2362	1487	1106
		0.05	−436	−517	368	−1376	−501	2608	1565	718
		0.01	13438	−586	627	−3195	−2013	16188	2202	−326
400	0.20	0.2	−3552	945	1067	915	1037	−1935	1337	1307
		0.1	−876	1258	1443	1143	1327	622	1702	1592
		0.05	978	452	686	207	440	3720	2247	2007
		0.01	10610	2763	3091	2023	2348	10536	2545	1854
			$\lambda = 3$							
4	0.33	0.2	−60592	870	11055	798	10957	−80114	−26533	−26533
		0.1	−57444	−5297	11281	−6718	9394	−29789	9855	7324
		0.05	−46634	−9869	11157	−14902	4613	−14678	10056	2930
		0.01	26638	−4060	24380	−31644	−9665	13370	1754	−20105
10	0.33	0.2	−39441	−2804	1303	−2993	1091	−25995	3221	2934
		0.1	−23805	987	7968	−893	5877	−16009	4545	2832
		0.05	−15414	−3773	4891	−8413	−248	−5879	4125	−81
		0.01	35693	−190	11690	−18443	−8302	21448	2930	−10368
40	0.33	0.2	−17080	−802	382	−969	210	−12693	712	568
		0.1	−8898	263	2131	−577	1268	−7330	1116	479
		0.05	−507	52	2439	−1940	394	−768	1687	216
		0.01	20786	−1625	1649	−7810	−4694	16434	2976	−1538
100	0.33	0.2	−8737	889	1399	790	1298	−7519	264	190
		0.1	−5590	−523	253	−922	−151	−4346	154	−142
		0.05	507	−110	887	−1028	−42	271	830	165
		0.01	12495	−2490	−1115	−5201	−3854	10512	971	−990
400	0.33	0.2	−4023	553	687	522	655	−1863	1108	1084
		0.1	−876	1179	1384	1057	1262	723	1901	1812
		0.05	2864	1947	2211	1679	1942	3545	2650	2456
		0.01	7781	−394	−29	−1150	−786	8637	2544	1995

Table 2. Values of $e_\alpha \times 10^5$ under hypothesis 1 for $\lambda = 1$ and $\lambda = 0.5$.

k	Q	α	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
			$\lambda = 1$							
		0.2	−13451	43890	64752	42194	62666	2640	43561	42689
4	1.00	0.1	−57444	−103	20759	−1799	18673	−41353	−432	−1304
		0.05	−93774	−36433	−15571	−38129	−17657	−77683	−36762	−37634
		0.01	50208	−6557	58406	−39642	14355	48472	42305	25160
		0.2	−39441	1328	5909	1092	5650	−36947	−10610	−10679
10	1.00	0.1	−38712	−8196	337	−9715	−1335	−5202	13714	12691
		0.05	−30321	−14668	−2651	−18875	−7290	−22968	−6677	−8177
		0.01	35693	−11609	8937	−31656	−13255	13207	9962	3558
		0.2	−17080	702	2308	492	2092	−9036	3329	3258
40	1.00	0.1	−16352	−6002	−3435	−6871	−4327	−5878	3597	3308
		0.05	−7960	−7914	−4459	−10057	−6656	−4339	1948	1326
		0.01	35693	−82	5341	−8497	−3286	8260	5415	3482
		0.2	−13451	−2625	−1944	−2730	−2050	−6446	1174	1142
100	1.00	0.1	−876	3024	4207	2469	3646	−3734	1964	1836
		0.05	5221	1764	3302	534	2060	−1203	2280	1998
		0.01	21923	−1068	1131	−4718	−2555	5668	3394	2566
		0.2	−6380	−1373	−1183	−1410	−1220	−1154	1471	1461
400	1.00	0.1	−876	790	1095	636	941	417	2027	1990
		0.05	5221	2857	3259	2505	2906	1597	2060	1981
		0.01	7781	−3527	−2971	−4473	−3918	4546	2088	1865
			$\lambda = 0.5$							
		0.2	−13451	51969	88598	49944	86042	17512	57260	55730
4	2.00	0.1	−57444	7976	44605	5951	42049	−26481	13267	11737
		0.05	−93774	−28354	8275	−30379	5719	−62811	−23064	−24593
		0.01	−161924	−96504	−59876	−98529	−62431	−130961	−91214	−92743
		0.2	−106523	−55965	−61417	−55907	−61351	−99554	−66045	−66041
10	2.00	0.1	−61073	−19806	−10687	−20691	−11709	−26112	−2538	−3252
		0.05	−7960	−3559	15857	−10554	7711	−18117	−4214	−6103
		0.01	13333	−33481	−6168	−51443	−27246	31169	8255	−201
		0.2	−39441	−16121	−14622	−16206	−14711	−34066	−17639	−17665
40	2.00	0.1	6009	10281	14552	8511	12714	−19341	−7734	−7979
		0.05	14400	3946	9405	246	5564	3047	6748	5920
		0.01	35693	−11994	−4456	−21847	−14688	15777	4152	1715
		0.2	−13451	−1328	−351	−1464	−489	−7410	1226	1188
100	2.00	0.1	−876	2211	3913	1502	3194	−250	4905	4739
		0.05	−8922	−11936	−9894	−13162	−11139	557	2196	1851
		0.01	36066	−188	3077	−5330	−2142	6679	−1037	−2022
		0.2	691	5697	6000	5635	5938	−2473	258	247
400	2.00	0.1	−876	234	682	34	481	−1394	−281	−324
		0.05	5221	1046	1634	590	1176	3036	2113	2017
		0.01	21923	3347	4201	1930	2778	7065	1250	976

Table 3. Values of $e_\alpha \times 10^5$ under hypothesis 2 for $\lambda = 5$ and $\lambda = 3$.

k	Q	α	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
			$\lambda = 5$							
4	0.31	0.2	−69319	−5074	6382	−5095	6350	−59186	−4958	−5020
		0.1	−56474	−3568	12384	−5185	10124	−40406	3028	1116
		0.05	−32037	−1556	15280	−8669	5930	−30911	−402	−5861
		0.01	33243	−6006	7162	−35994	−28238	22432	5586	−18021
10	0.37	0.2	−38471	−98	4161	−341	3881	−31623	−339	−555
		0.1	−28679	−2473	3091	−4427	868	−22780	62	−1392
		0.05	−10808	−2842	3252	−8597	−3203	−11118	566	−3296
		0.01	43407	−4566	1116	−25936	−22216	20906	1245	−11930
40	0.41	0.2	−17013	44	1068	−172	844	−13555	505	365
		0.1	−9200	−782	638	−1835	−451	−7472	1242	606
		0.05	2157	−617	1061	−3202	−1609	−1691	762	−689
		0.01	34382	−96	1873	−8615	−6897	17789	2780	−1798
100	0.42	0.2	−10320	−196	220	−318	96	−7771	−1404	−1477
		0.1	−4888	−690	−93	−1215	−626	−3574	−803	−1105
		0.05	2263	−796	−68	−2009	−1298	609	−805	−1475
		0.01	20651	−1798	−874	−5524	−4649	10499	−1684	−3649
400	0.43	0.2	−4357	347	455	306	414	−2427	−630	−653
		0.1	−1031	487	647	327	486	337	216	127
		0.05	3255	985	1185	629	829	2844	529	337
		0.01	14784	2455	2726	1387	1654	9924	1881	1322
			$\lambda = 3$							
4	0.51	0.2	−77553	−9315	998	−9317	994	−61733	−7980	−8020
		0.1	−56368	−1359	14798	−3154	12309	−45049	−745	−2252
		0.05	−41308	−6611	10389	−13113	1861	−36024	−3657	−8203
		0.01	34992	−11142	2064	−42165	−33891	8992	−276	−19907
10	0.61	0.2	−41134	−108	4151	−346	3876	−32003	843	649
		0.1	−27649	−1195	4413	−3465	1839	−22952	1666	327
		0.05	−9542	−3706	2314	−10203	−4939	−11878	1993	−1551
		0.01	60956	−4538	630	−30170	−27043	20454	3515	−8787
40	0.68	0.2	−18511	−408	624	−648	376	−13382	747	617
		0.1	−8802	−952	478	−2193	−804	−8060	940	357
		0.05	3688	−1615	53	−4667	−3097	−582	2037	662
		0.01	40346	−2818	−924	−12930	−11318	15390	1605	−2568
100	0.70	0.2	−10790	−203	221	−346	75	−8501	−504	−570
		0.1	−4809	−1157	−553	−1781	−1185	−4252	255	−21
		0.05	3616	−1382	−648	−2839	−2125	170	578	−41
		0.01	29176	−296	626	−4996	−4132	12736	2201	323
400	0.71	0.2	−4864	11	121	−38	72	−2474	586	564
		0.1	−1142	−8	156	−200	−37	−149	1066	984
		0.05	3696	355	560	−73	130	1965	1070	895
		0.01	15605	509	783	−770	−500	6568	319	−181

Table 4. Values of $e_\alpha \times 10^5$ under hypothesis 2 for $\lambda = 1$ and $\lambda = 0.5$.

k	Q	α	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
			$\lambda = 1$							
		0.2	−80978	841	8740	840	8739	−58847	−2808	−2853
4	1.53	0.1	−66143	1509	16723	−183	14382	−65207	−14087	−14669
		0.05	16142	16997	33598	502	13640	−35502	−197	−3994
		0.01	−2137	−36517	−20894	−61308	−49617	45954	23780	1048
		0.2	−49410	1305	5024	1101	4790	−34199	−514	−661
10	1.83	0.1	−36350	−3765	1560	−6316	−1306	−25550	203	−878
		0.05	−9162	−7664	−1816	−16116	−11117	−13849	1118	−1863
		0.01	76392	−22622	−17622	−53446	−50197	19992	3462	−7326
		0.2	−22260	−553	457	−878	121	−14935	430	323
40	2.05	0.1	−10307	−3295	−1854	−5114	−3728	−9203	1037	538
		0.05	8213	−5180	−3489	−9907	−8347	−829	2880	1669
		0.01	65091	−10406	−8535	−26689	−25183	15631	2653	−1070
		0.2	−12677	−435	1	−653	−220	−7619	2147	2087
100	2.11	0.1	−3456	−1772	−1134	−2787	−2162	−3645	2621	2374
		0.05	8277	−3323	−2549	−5732	−4987	307	2494	1946
		0.01	47886	−3810	−2846	−11912	−11036	11901	3213	1561
		0.2	−5758	−354	−234	−433	−314	−2475	756	736
400	2.14	0.1	129	−197	−18	−522	−344	−858	565	494
		0.05	7126	−117	108	−853	−630	1294	616	464
		0.01	28078	1549	1852	−763	−467	6946	834	391
			$\lambda = 0.5$							
		0.2	−50802	37493	42888	37035	42258	−83664	−27627	−27627
4	3.06	0.1	−70549	8225	17263	6451	14862	−34938	8844	7367
		0.05	−68571	−10261	2700	−15734	−4536	−30352	2253	−1500
		0.01	−38166	−48921	−30760	−68163	−55065	−13441	−11608	−23765
		0.2	−45567	10817	13042	10448	12628	−31536	1642	1483
10	3.65	0.1	−41248	−3200	970	−6038	−2177	−21599	3049	1930
		0.05	−17665	−12548	−7023	−21434	−16753	−9672	3697	697
		0.01	80515	−34638	−27775	−67073	−62368	25785	6001	−4720
		0.2	−25533	−723	120	−1112	−280	−14358	1474	1367
40	4.11	0.1	−14220	−6350	−4955	−8572	−7234	−8398	2006	1506
		0.05	6062	−10499	−8713	−16376	−14729	−1768	2059	889
		0.01	82635	−17338	−15004	−38950	−37039	17755	3410	−323
		0.2	−15006	−1252	−840	−1533	−1125	−8440	1544	1486
100	4.21	0.1	−4538	−3738	−3079	−5100	−4456	−4247	2044	1802
		0.05	11536	−5095	−4249	−8479	−7667	1295	3096	2542
		0.01	59148	−9002	−7865	−20319	−19284	14489	4469	2782
		0.2	−6380	−577	−453	−688	−565	−2538	1963	1944
400	4.27	0.1	289	−1086	−893	−1549	−1358	−499	2059	1988
		0.05	8583	−1559	−1313	−2616	−2372	2517	2780	2624
		0.01	32776	−1840	−1497	−5158	−4823	7484	2000	1554

Table 5. Values of $e_\alpha \times 10^5$ under hypothesis 3 for $\lambda = 5$ and $\lambda = 3$.

k	Q	α	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
			$\lambda = 5$							
4	2.05	0.2	−89388	−1920	9167	−1917	9171	−104197	−32855	−32794
		0.1	−86345	−7036	5504	−7906	4424	−80600	−13631	−14132
		0.05	55019	23965	27282	−1661	2261	−51161	−5098	−9640
		0.01	70489	−27370	−28625	−66417	−64209	3210	−11761	−33905
10	2.05	0.2	−52761	−713	2585	−906	2372	−43760	−1470	−1613
		0.1	−17424	5938	8855	1207	3793	−30349	214	−1375
		0.05	−4174	−7435	−5346	−17934	−16383	−12604	828	−4001
		0.01	145832	−10764	−13338	−56519	−58708	33040	−2795	−20287
40	2.05	0.2	−21360	−412	217	−797	−178	−17914	577	415
		0.1	−5856	−1089	−411	−3347	−2719	−10611	326	−485
		0.05	14006	−2774	−2177	−8506	−8020	−1623	−125	−2086
		0.01	68156	−10115	−9934	−28528	−28599	25007	−163	−6590
100	2.05	0.2	−10994	525	769	265	505	−10898	89	−4
		0.1	−634	−144	152	−1346	−1061	−5322	473	68
		0.05	10073	−2602	−2293	−5348	−5064	2394	1704	751
		0.01	49584	−3240	−2983	−12304	−12116	18671	1200	−1706
400	2.05	0.2	−4723	336	398	245	306	−3232	−963	−996
		0.1	412	−211	−127	−573	−490	−847	−1299	−1422
		0.05	6937	−524	−424	−1335	−1238	2882	−816	−1087
		0.01	25543	−548	−430	−3030	−2918	9999	−1907	−2688
			$\lambda = 3$							
4	3.42	0.2	−59906	28783	37563	28522	37251	−84391	−13922	−13922
		0.1	−88738	−4829	3983	−5741	2915	−65696	−4260	−5392
		0.05	−56847	−8222	−1102	−16744	−10298	−35042	3437	−3329
		0.01	200950	−12512	−16572	−69199	−67179	57845	13034	−22734
10	3.42	0.2	−55397	−318	2096	−507	1892	−44180	−2475	−2629
		0.1	−46655	−9062	−6600	−11827	−9528	−31177	−627	−2334
		0.05	32074	4551	5771	−13055	−12308	−11482	2014	−3368
		0.01	102953	−25937	−26730	−67150	−67905	38360	2574	−17361
40	3.42	0.2	−21839	−91	402	−562	−78	−17207	1015	829
		0.1	−5905	−1851	−1304	−4598	−4096	−10143	763	−141
		0.05	12897	−5700	−5200	−12440	−12035	−276	1300	−905
		0.01	88571	−7947	−7790	−32038	−32101	23098	−73	−6989
100	3.42	0.2	−12909	−833	−639	−1138	−947	−10056	618	512
		0.1	−2324	−2227	−1985	−3664	−3433	−5639	104	−340
		0.05	12891	−2737	−2480	−6238	−6004	986	638	−389
		0.01	53895	−5923	−5698	−17121	−16959	17398	908	−2246
400	3.42	0.2	−5328	−194	−144	−308	−258	−2931	1670	1633
		0.1	1466	73	144	−395	−326	−801	1163	1027
		0.05	7585	−1317	−1234	−2346	−2265	2800	1621	1321
		0.01	30397	−124	−24	−3355	−3261	12279	2851	1964

Table 6. Values of $e_\alpha \times 10^5$ under hypothesis 3 for $\lambda = 1$ and $\lambda = 0.5$.

k	Q	α	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
			$\lambda = 1$							
4	10.26	0.2	−81105	19116	16043	19115	16042	−18462	44840	43262
		0.1	−125098	−24877	−27950	−24878	−27951	−12346	30993	26392
		0.05	52562	26053	26963	−7073	−7630	−48676	−5337	−9938
		0.01	14982	−36026	−34812	−74639	−75159	−2307	972	−21220
10	10.26	0.2	7675	42995	43435	37519	37858	−30107	5048	4755
		0.1	−16984	8196	8775	−4	426	−16043	9722	7617
		0.05	−24312	−15876	−15117	−28841	−28313	−5312	9570	4420
		0.01	62211	−37264	−35867	−80455	−79749	9101	−427	−14525
40	10.26	0.2	−21326	1644	1811	839	1001	−16133	−337	−495
		0.1	−16984	−9417	−9122	−12855	−12578	−9712	951	196
		0.05	4689	−12725	−12310	−21868	−21498	−2643	1835	64
		0.01	110546	−7274	−6617	−44775	−44284	15099	3273	−2179
100	10.26	0.2	−16908	−3570	−3481	−4014	−3927	−9063	1782	1693
		0.1	−2818	−3804	−3645	−6137	−5983	−4860	2505	2136
		0.05	18935	−2104	−1887	−8131	−7927	221	3443	2609
		0.01	70008	−5013	−4703	−23942	−23669	13169	5525	3014
400	10.26	0.2	−4680	395	426	191	222	−3276	−436	−465
		0.1	1768	−800	−751	−1608	−1560	−1092	−97	−207
		0.05	11293	−772	−709	−2609	−2547	1078	−41	−277
		0.01	37910	−286	−195	−5955	−5867	5581	−870	−1540
			$\lambda = 0.5$							
4	20.51	0.2	−105109	−989	−12486	343	−11275	−73325	−9264	−9270
		0.1	−149102	−44982	−56479	−43651	−55268	−117318	−53257	−53263
		0.05	−185432	−81312	−92809	−79981	−91598	−153648	−89587	−89593
		0.01	154038	−6232	−1978	−71892	−76552	5577	2321	−20234
10	20.51	0.2	−95799	−26872	−29541	−26814	−29484	−54535	−16492	−16542
		0.1	75043	45288	46058	15050	15231	−10987	13364	11121
		0.05	53035	13185	14165	−20263	−19987	−16038	585	−3525
		0.01	13530	−47038	−45650	−86854	−86421	5117	−6253	−19143
40	20.51	0.2	−21502	3726	3676	2565	2516	−16851	212	67
		0.1	−29689	−16224	−16127	−19811	−19714	−6344	4686	3914
		0.05	19915	−5765	−5322	−20406	−19987	1978	5998	4186
		0.01	94988	−16010	−15033	−58957	−58145	20488	6418	943
100	20.51	0.2	−16792	−2303	−2273	−2966	−2935	−7965	1318	1230
		0.1	2623	−385	−252	−4092	−3958	−3935	1597	1240
		0.05	11584	−8192	−7985	−15787	−15582	2105	3045	2230
		0.01	83836	−3177	−2750	−30537	−30131	13898	3245	844
400	20.51	0.2	−6035	−459	−439	−754	−733	−3239	988	961
		0.1	2057	−1182	−1141	−2390	−2349	−1046	1208	1105
		0.05	13283	−929	−868	−3698	−3637	1752	1700	1475
		0.01	42508	−1287	−1188	−9687	−9588	8034	2085	1434

Table 7. Minimum values of k that satisfy the condition $e_{\max} < 0.1$ under hypothesis 1.

λ	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
5	—	*10	*10	40	40	—	*10	40
3	400	*4	40	40	10	400	10	40
1	400	40	*10	100	40	40	40	40
0.5	—	400	*100	400	400	*100	*100	*100

Table 8. Minimum values of k that satisfy the condition $e_{\max} < 0.1$ under hypothesis 2.

λ	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
5	—	*4	10	40	40	400	*4	40
3	—	10	10	100	100	400	*4	10
1	—	100	40	400	400	400	*10	*10
0.5	—	100	100	400	400	400	*10	*10

Table 9. Minimum values of k that satisfy the condition $e_{\max} < 0.1$ under hypothesis 3.

λ	X	g_1X	$g_{1E}X$	g_2X	$g_{2E}X$	G	g_1G	g_2G
5	—	100	40	400	400	400	*10	40
3	—	40	40	400	400	—	*10	40
1	—	100	100	400	400	400	*10	40
0.5	—	100	100	400	400	400	*40	*40

of k that satisfy the condition $e_{\max} < 0.1$ are shown in Tables 7–9. In these tables, the minimum value is indicated by the symbol * for each λ. From Tables 1, 2 and 7, we obtain the following results under hypothesis 1.

- The performance of the transformation g_1 and g_{1E} is better than that of the transformation g_2 and g_{2E}.
- $g_1(S_k)$, $g_1(T_k)$, and $g_{1E}(T_k)$ perform well for $k \geq 10$ when $\lambda = 5$ and 3, for $k \geq 40$ when $\lambda = 1$, and for $k \geq 100$ when $\lambda = 0.5$.

From Tables 3, 4 and 8, we obtain the following results under hypothesis 2.

- The performance of the transformation g_1 and g_{1E} is better than that of the transformation g_2 and g_{2E}.
- $g_1(S_k)$ performs well for $k \geq 4$ when $\lambda = 5$ and 3 and for $k \geq 10$ when $\lambda = 1$ and 0.5.
- $g_1(T_k)$ and $g_{1E}(T_k)$ perform well for $k \geq 10$ when $\lambda = 5$ and 3, for $k \geq 40$ when $\lambda = 1$, and for $k \geq 100$ when $\lambda = 0.5$.

From Tables 5, 6 and 9, we obtain the following results under hypothesis 3.

- The performance of the transformation g_1 and g_{1E} is better than that of the transformation g_2 and g_{2E}.
- $g_1(S_k)$ performs well for $k \geq 10$ when $\lambda = 5$, 3, and 1 and for $k \geq 40$ when $\lambda = 0.5$.
- $g_1(T_k)$ and $g_{1E}(T_k)$ perform well for $k \geq 40$ when $\lambda = 5$ and 3 and for $k \geq 100$ when $\lambda = 1$ and 0.5.

From Tables 1–9, we obtain the following combined results.

- The performance of $g_1(T_k)$ and $g_{1E}(T_k)$ is better than that of $g_2(T_k)$ and $g_{2E}(T_k)$.

- $g_1(T_k)$ and $g_{1E}(T_k)$ improve the standardized X_k^2 statistic.
- The performance of $g_1(S_k)$ is better than that of $g_2(S_k)$.
- $g_1(S_k)$ and $g_2(S_k)$ improve the standardized G_k^2 statistic.

The above results indicate that the performance of the transformation g_1 is better than that of g_2. We therefore evaluate the performance of the transformation g_1 more precisely. From Tables 1–9 and numerous other examples, we obtain the following further results regarding g_1.

- $g_1(S_k)$ is very accurate for $k \geq 10$ when $\lambda > 1$ to such a degree that it satisfies the condition $e_{\max} < 0.1$ for all hypotheses and situations.
- The performance of $g_1(S_k)$ is better on the whole than that of $g_1(T_k)$ and $g_{1E}(T_k)$.
- $g_1(T_k)$ and $g_{1E}(T_k)$ are very accurate for $k \geq 40$ when $\lambda > 1$ to such a degree that they satisfy the condition $e_{\max} < 0.1$ for almost all hypotheses and situations.
- As for comparison of the performance of $g_1(T_k)$ and that of $g_{1E}(T_k)$, $g_1(T_k)$ performs better than does $g_{1E}(T_k)$ when $Q \leq 0.5$, while $g_{1E}(T_k)$ performs better than does $g_1(T_k)$ when $Q > 1$.

6. Concluding remarks

We have proposed new goodness-of-fit statistics for sparse multinomials. These are improvements in asymptotic normality based on normalizing transformations.

The results of numerical investigations presented in Section 5 indicate that the transformation g_1 greatly improves the asymptotic normality of the test statistics. The performance of $g_1(S_k)$ is better than that of the others. However, $g_1(S_k)$ requires simulation for estimating μ_{S1} defined by (4.9). Therefore, we recommend $g_1(S_k)$ in a situation in which μ_{S1}^* defined by (4.11) can be calculated, and we recommend $g_1(T_k)$ and $g_{1E}(T_k)$ for other situations.

In future works, we will study the power of the statistics to investigate their performance more precisely. Koehler (1979) derived a general method for obtaining exact moments of X_k^2. Since this is applicable to alternative cases, our method can be extended to alternative cases by using the exact moments.

Appendix A: Reason for substituting ν_S' for ν_S

If we define

$$f_{jk}(y) = 2I(y, n_k p_{0jk}) - 2E[I(Y_{jk}, n_k p_{0jk})] - \gamma_k(y - n_k p_{0jk}), \tag{A.1}$$

the log-likelihood ratio statistic G_k^2 is represented as

$$G_k^2 = \sum_{j=1}^{k} f_{jk}(X_{jk}) + \mu_{LR,k}, \tag{A.2}$$

where $\mu_{LR,k}$ is given by (2.4). Thus, by the definition of S_k, (4.5), and (A.2), we have

$$\nu_S = \lim_{k\to\infty} \frac{1}{k} E\left[\left\{\sum_{j=1}^{k} f_{jk}(X_{jk}) - E\left(\sum_{j=1}^{k} f_{jk}(X_{jk})\right)\right\}^3\right]. \tag{A.3}$$

Next, we consider an expression constructed by substituting Y_{jk} for X_{jk} in the right-hand side of expression (A.3). By the definition of Y_{jk}, (A.1), (4.8), and (4.7), we have

$$\lim_{k\to\infty}\frac{1}{k}E\left[\left\{\sum_{j=1}^{k} f_{jk}(Y_{jk}) - E\left(\sum_{j=1}^{k} f_{jk}(Y_{jk})\right)\right\}^3\right] = \lim_{k\to\infty}\frac{\nu_{LR,k}}{k} = \nu'_S. \tag{A.4}$$

Here, we discuss the relation between the limit value derived in the right-hand side of (A.3) and that derived in the left-hand side of (A.4). By Theorem 5.2 and its proof given by Morris (1975), both the statistic $s_{LR,k}^{-1}\sum_{j=1}^{k} f_{jk}(X_{jk})$ and the statistic $s_{LR,k}^{-1}\sum_{j=1}^{k} f_{jk}(Y_{jk})$ are asymptotically distributed according to the standard normal distribution as k tends to infinity, where $s_{LR,k}$ is defined by (2.5). If these two statistics satisfy the condition of uniform integrability, the limit value derived in the right-hand side of (A.3) coincides with that derived in the left-hand side of (A.4). Though we cannot prove uniform integrability of the two statistics, the two limit values are considered to be very similar. Therefore, we substitute ν'_S for ν_S.

Appendix B: Proof of $|\mu_S| < \infty$, $\sigma_S^2 < \infty$, and $|\nu'_S| < \infty$

Let Y be distributed as $Po(m)$, where $Po(m)$ denotes a Poisson random variable with mean m. We can prove that the following inequalities hold by using the Schwartz inequality.

$$0 \le E[I(Y,m)] \le 1. \tag{B.1}$$

$$0 \le V[I(Y,m)] \le 1 + \frac{1}{m}. \tag{B.2}$$

$$-\left(\frac{m}{2}+1\right) \le \mathrm{Cov}[I(Y,m),Y] \le \frac{m}{2}+1. \tag{B.3}$$

$$-1 \le E[\{I(Y,m) - E[I(Y,m)]\}^3] \le \frac{1}{m^2}. \tag{B.4}$$

$$-m \le E[\{I(Y,m) - E[I(Y,m)]\}(Y-m)^2] \le 1. \tag{B.5}$$

$$-\left(\frac{1}{2}+m+\frac{1}{m}\right) \le E[\{I(Y,m) - E[I(Y,m)]\}^2(Y-m)] \le \frac{1}{2}+m+\frac{1}{m}. \tag{B.6}$$

By applying (B.1)–(B.6) to the case in which Y_{jk}, $(j = 1,\ldots,k)$ is distributed as $Po(m_{jk})$, $(j = 1,\ldots,k)$, where $m_{jk} = n_k p_{0jk}$, we find that

$$0 \le \lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k} E[I(Y_{jk},m_{jk})] \le 1, \tag{B.7}$$

$$0 \le \lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k} V[I(Y_{jk},m_{jk})] \le 1 + d, \tag{B.8}$$

$$-\left(1+\frac{2}{c}\right) \le \lim_{k\to\infty}\gamma_k = \lim_{k\to\infty}\frac{2}{n_k}\sum_{j=1}^{k}\mathrm{Cov}[I(Y_{jk},m_{jk}),Y_{jk}] \le 1+\frac{2}{c}, \tag{B.9}$$

$$-1 \le \lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k} E[\{I(Y_{jk},m_{jk}) - E[I(Y_{jk},m_{jk})]\}^3] \le e, \tag{B.10}$$

$$\text{(B.11)} \qquad -c \leq \lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} E[\{I(Y_{jk}, m_{jk}) - E[I(Y_{jk}, m_{jk})]\}(Y_{jk} - m_{jk})^2] \leq 1,$$

and

$$\text{(B.12)} \qquad \lim_{k\to\infty} \left| \frac{1}{k} \sum_{j=1}^{k} E[\{I(Y_{jk}, m_{jk}) - E[I(Y_{jk}, m_{jk})]\}^2 (Y_{jk} - m_{jk})] \right| \leq \frac{1}{2} + c + d,$$

where c, d, and e are given by (3.1), (3.5), and (3.6), respectively. Therefore, $|\mu_S| < \infty$ is proved by (B.7), $\sigma_S^2 < \infty$ is proved by (B.8) and (B.9), and $|\nu'_S| < \infty$ is proved by (B.9)–(B.12).

Acknowledgements

The authors are very grateful to the referees for their valuable comments and suggestions. This research was supported in part by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (No. 11308010).

References

Haldane, J. B. S. (1937). The exact value of the moments of the distribution of χ^2, used as a test of goodness of fit, when expectations are small, *Biometrika*, **29**, 133–143.

Hall, P. (1992). On the removal of skewness by transformation, *J. Roy. Statist. Soc. Ser. B*, **54**, 221–228.

Holst, L. (1972). Asymptotic normality and efficiency for certain goodness-of-fit tests, *Biometrika*, **59**, 137–145.

Koehler, K. J. (1979). A general formula for moments of the pearson goodness-of-fit statistic for alternatives, *Biometrika*, **66**, 397–399.

Koehler, K. J. (1986). Goodness-of-fit tests for log-linear models in sparse contingency tables, *J. Amer. Statist. Assoc.*, **81**, 483–493.

Koehler, K. J. and Larntz, K. (1980). An empirical investigation of goodness-of-fit statistics for sparse multinomials, *J. Amer. Statist. Assoc.*, **75**, 336–344.

Konishi, S. (1981). Normalizing transformations of some statistics in multivariate analysis, *Biometrika*, **68**, 647–651.

Konishi, S. (1991). Normalizing transformations and bootstrap confidence intervals, *Ann. Statist.*, **19**, 2209–2225.

Morris, C. (1975). Central limit theorems for multinomial sums, *Ann. Statist.*, **3**, 165–188.

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.

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., Sekiya, Y. and Suzukawa, A. (2002*a*). Asymptotic approximations for the distributions of the multinomial goodness-of-fit statistics under local alternatives, *J. Multivariate Anal.*, **81**, 335–359.

Taneichi, N., Sekiya, Y., Suzukawa, A. and Imai, H. (2002*b*). A formula for normalizing transformation of some statistics, *Comm. Statist. Theory Methods*, **31**, 163–179.

Yanagihara, H. and Tonda, T. (2003). Adjustment on an asymptotic expansion of the distribution function with χ^2-approximation, *Hiroshima Math. J.*, **33**, 15–25.

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.