正規化変換と検定統計量への応用

(Normalizing Transformations and Their Applications to Test Statistics)

帯広畜産大学畜産学部 種市信裕
北海道教育大学教育学部 関谷祐里
帯広畜産大学畜産学部 鈴川晶夫
北海道大学大学院工学研究科 今井英幸

Abstract

On the basis of Konishi's discussion of finding a normalizing transformation (1981), two types of concrete normalizing transformations are derived. The proposed normalizing transformations are applied to functions of a sample covariance matrix. Performance of the transformations in the applications is numerically compared.

1 Introduction

Normalizing transformations are useful for obtaining simple and accurate approximations to the distributions of statistics. Especially, they are considered to be effective for the statistic which is very difficult to find a valid asymptotic expansion. The statistic which is composed by discrete random variables and the statistic which includes nuisance parameters in its formal expansion are the typical examples. From now on, we consider statistics whose limiting distributions are normal. Konishi (1981) discussed a general method to find a normalizing transformation when we have already obtained the asymptotic expansion for the transformed statistic. In Konishi (1981), it was shown that we find a normalizing transformation as a solution of a differential equation. Konishi (1991) also derived the differential equation for making a normalizing transformation under certain assumptions of moments. In this article, we propose two types of concrete normalizing transformations. One transformation is represented as a power function. Another transformation is represented as a exponential function. In Section 2, the normalizing transformations are proposed. In Section 3, they are applied to functions of a sample covariance matrix. In Section 4, performance of the transformations in the application is numerically compared.

2 Normalizing transformations of a random variable under certain assumptions of moments

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 are evaluated as

$$E(T_n) = \mu + \frac{1}{n}\mu_1 + o\left(\frac{1}{n}\right), \tag{2.1}$$

$$V(T_n) = \frac{1}{n}\sigma^2 + o\left(\frac{1}{n}\right), \tag{2.2}$$

and

$$E[\{T_n - E(T_n)\}^3] = \frac{1}{n^2}\nu + o\left(\frac{1}{n^2}\right), \tag{2.3}$$

respectively, where $\sigma \neq 0$ and $\nu \neq 0$. We also assume that the distribution function of

$$\frac{\sqrt{n}(T_n - \mu)}{\sigma}$$

tends to a standard normal distribution function as $n \to \infty$. Let $f(x)$ be a strictly monotone and twice continuously differentiable function in a neighbourhood of $x = \mu$. If we define

$$g(x) = \frac{\sqrt{n}}{f'(\mu)\sigma}\left\{f(x) - f(\mu) - \frac{1}{n}\left(f'(\mu)\mu_1 + \frac{1}{2}f''(\mu)\sigma^2\right)\right\}, \tag{2.4}$$

$g(x)$ is also a strictly monotone and twice continuously differentiable function in a neighbourhood of $x = \mu$. If we denote the cumulant-generating function of $g(T_n)$ by $\psi_{g(T_n)}(t)$, it is evaluated as

$$\psi_{g(T_n)}(t) = \frac{(it)^2}{2} + \frac{1}{\sqrt{n}}\frac{(it)^3}{6\sigma^3}\left(\nu + 3\frac{f''(\mu)}{f'(\mu)}\sigma^4\right) + o\left(\frac{1}{\sqrt{n}}\right). \tag{2.5}$$

We consider a case that T_n is a continuous random variable which satisfies certain regularity conditions. By applying the inversion formula to (2.5), $F(x) = P(g(T_n) < x)$ is evaluated as

$$F(x) = \Phi(x) - \frac{1}{\sqrt{n}}\frac{1}{6\sigma^3}\left(\nu + 3\frac{f''(\mu)}{f'(\mu)}\sigma^4\right)(x^2 - 1)\phi(x) + o\left(\frac{1}{\sqrt{n}}\right),$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution function and its first derivative, respectively. Therefore, if f satisfies the condition

$$\nu + 3\frac{f''(\mu)}{f'(\mu)}\sigma^4 = 0, \tag{2.6}$$

g is a normalizing transformation in the sense that

$$F(x) = \Phi(x) + o\left(\frac{1}{\sqrt{n}}\right). \tag{2.7}$$

On the other hand, we consider a case that T_n is a discrete random variable. If f satisfies the condition (2.6), g is a normalizing transformation in the sense that

$$g(T_n) \xrightarrow{L} N(0,1), \text{ as } n \to \infty$$

and

$$E[\{g(T_n) - E(g(T_n))\}^3] = o\left(\frac{1}{\sqrt{n}}\right). \tag{2.8}$$

The condition (2.6) which is necessary to obtain a normalizing transformation under assumptions (2.1), (2.2), and (2.3) was derived in Konishi (1991).

In this paper, we derive concrete functions which satisfy the condition (2.6). When $\mu \neq 0$, we consider the function

$$f_1(x) = \begin{cases} \frac{\mu}{\eta}\left\{\left(\frac{x}{\mu}\right)^{\eta} - 1\right\} & (\eta \neq 0) \\ \mu \log \frac{x}{\mu} & (\eta = 0). \end{cases} \tag{2.9}$$

For any η, let the domain of f_1 be $(0, \infty)$ in the case of $\mu > 0$, while let the domain of f_1 be $(-\infty, 0)$ in the case of $\mu < 0$. On the other hand, for $\eta > 0$ we may adopt $[0, \infty)$ or $(-\infty, 0]$ as domain. Here, if we put

$$\eta = -\frac{\mu\nu}{3\sigma^4} + 1, \tag{2.10}$$

then f_1 is not only strictly monotone and twice continuously differentiable in a neighbourhood of $x = \mu$, but it also satisfies the condition (2.6). Therefore, if we apply f_1 to (2.4), then

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

where

$$\xi = -\frac{\nu}{3\sigma^4}, \tag{2.11}$$

is a normalizing transformation of T_n which satisfies (2.7) or (2.8). In the case of $\mu \neq 0$, a concrete normalizing transformation of T_n is derived by g_1.

Next, we derive another transformation. When $\sigma \neq 0$ and $\nu \neq 0$, we consider the function

$$f_2(x) = \frac{1}{\xi} e^{\xi(x-\mu)} - \frac{1}{\xi}, \tag{2.12}$$

where ξ is given by (2.11). Then, f_2 is not only an infinitely differentiable and strictly monotone function in domain$(-\infty, \infty)$, but it also satisfies the condition (2.6). If we apply f_2 to (2.4), then

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

is a normalizing transformation of T_n which satisfies (2.7) or (2.8). Even when $\mu = 0$, we can apply g_2 directly.

3 Examples

3.1 Functions of a sample covariance matrix

Let $\boldsymbol{S} = (s_{ij})$ be the sample (unbiased) covariance matrix based on a random sample of size $N = n + 1$ drawn from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a positive definite matrix. Let h be a real valued function which satisfies the following conditions (a) and (b).
(a) All the derivatives of $h(\boldsymbol{S})$ of order 3 and less are continuous in a neighbourhood of $\boldsymbol{S} = \boldsymbol{\Sigma}$.
(b) There exists at least one pair of (i, j) such that $\partial h(\boldsymbol{S})/\partial s_{ij}|_{\boldsymbol{S}=\boldsymbol{\Sigma}} \neq 0$.
Here, we consider the statistic

$$T_n = h(\boldsymbol{S}).$$

A stochastic approximation to T_n up to the order $n^{-3/2}$ is

$$T_n' = h(\boldsymbol{\Sigma}) + n^{-\frac{1}{2}}\mathrm{tr}(\boldsymbol{A}\boldsymbol{U}) + n^{-1}q_1(\boldsymbol{U}) + n^{-\frac{3}{2}}q_2(\boldsymbol{U}),$$

where

$$\boldsymbol{U} = (u_{ij}) = \sqrt{n}(\boldsymbol{S} - \boldsymbol{\Sigma}),$$

$$\boldsymbol{A} = \left\{\frac{1}{2}(1+\delta_{ij})\left.\frac{\partial h(\boldsymbol{S})}{\partial s_{ij}}\right|_{\boldsymbol{S}=\boldsymbol{\Sigma}}\right\},$$

$$q_1(\boldsymbol{U}) = \frac{1}{2}\sum_{i\geq j}\sum_{k\geq l} u_{ij}u_{kl}\left.\frac{\partial^2 h(\boldsymbol{S})}{\partial s_{ij}\partial s_{kl}}\right|_{\boldsymbol{S}=\boldsymbol{\Sigma}},$$

and

$$q_2(\boldsymbol{U}) = \frac{1}{6}\sum_{i\geq j}\sum_{k\geq l}\sum_{m\geq r} u_{ij}u_{kl}u_{mr}\left.\frac{\partial^3 h(\boldsymbol{S})}{\partial s_{ij}\partial s_{kl}\partial s_{mr}}\right|_{\boldsymbol{S}=\boldsymbol{\Sigma}}.$$

Then, the mean, variance, and third moment about the mean of T_n' are evaluated 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

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

where

$$\mu = h(\mathbf{\Sigma}),$$

$$\mu_1 = \frac{1}{2}\sum_{i\geq j}\sum_{k\geq l}(\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}) \left.\frac{\partial^2 h(\boldsymbol{S})}{\partial s_{ij}\partial s_{kl}}\right|_{\boldsymbol{S}=\mathbf{\Sigma}},$$

$$\sigma^2 = 2\mathrm{tr}(\boldsymbol{A}\mathbf{\Sigma})^2,$$

$$\nu = 8\mathrm{tr}(\boldsymbol{A}\mathbf{\Sigma})^3 + 12\sum_{i\geq j}\sum_{k\geq l}[\mathbf{\Sigma}\boldsymbol{A}\mathbf{\Sigma}]_{ij}[\mathbf{\Sigma}\boldsymbol{A}\mathbf{\Sigma}]_{kl} \left.\frac{\partial^2 h(\boldsymbol{S})}{\partial s_{ij}\partial s_{kl}}\right|_{\boldsymbol{S}=\mathbf{\Sigma}},$$

and $[\boldsymbol{A}]_{ij}$denotes the (i,j) element of a matrix $\boldsymbol{A}$ (Siotani et al., 1985, pp.161–162). Furthermore,

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

Then $g_1(T_n')$ and $g_2(T_n')$ are normalizing transformations of T_n' in the sense of (2.7).

3.1.1 The i-th largest characteristic root of a sample covariance matrix

The example given by Section 3.1 has many applications. As an application of the function of a sample covariance matrix, we pick up the i-th largest characteristic root of a sample covariance matrix. Let $\kappa_{in}(i = 1, \ldots, p)$ be the i-th largest characteristic root of the sample covariance matrix $\boldsymbol{S}$ based on a sample of size $N = n+1$ from a p-variate normal distribution with the population covariance matrix $\mathbf{\Sigma}$, and let $\lambda_1 \geq \ldots \geq \lambda_{i-1} > \lambda_i > \lambda_{i+1} \geq \ldots \geq \lambda_p$ be the ordered characteristic roots of $\mathbf{\Sigma}$. If we put $T_{in} = \kappa_{in}(i = 1, \ldots, p)$, under the assumption of simplicity, then the mean, variance, and the third moment about the mean of $T_{in}(i = 1, \ldots, p)$ are evaluated as $E(T_{in}) = \mu + n^{-1}\mu_1 + o(n^{-1})$, $V(T_{in}) = n^{-1}\sigma^2 + o(n^{-1})$, and $E[\{T_{in} - E(T_{in})\}^3] = n^{-2}\nu + o(n^{-2})$, where $\mu = \lambda_i$,

$$\mu_1 = \sum_{k\neq i}^{p}\lambda_{ik}\lambda_i\lambda_k,$$

$\sigma^2 = 2\lambda_i^2$, $\nu = 8\lambda_i^3$, and $\lambda_{jk} = (\lambda_j - \lambda_k)^{-1}$. Furthermore, under the assumption of simplicity,

$$\sqrt{n}\left(\frac{T_{in} - \mu}{\sigma}\right) \xrightarrow{L} N(0,1) \quad (i = 1, \ldots, p), \text{ as } n \to \infty.$$

Therefore,

$$g_1(T_{in}) = \sqrt{\frac{n}{2}}\left[3\left\{\left(\frac{T_{in}}{\lambda_i}\right)^{\frac{1}{3}} - 1\right\} - \frac{1}{n}\left(\sum_{k\neq i}^{p}\lambda_{ik}\lambda_k - \frac{2}{3}\right)\right]$$

and

$$g_2(T_{in}) = \sqrt{\frac{n}{2}}\left[-\frac{3}{2}\left\{e^{-\frac{2}{3}\left(\frac{T_{in}}{\lambda_i}-1\right)} - 1\right\} - \frac{1}{n}\left(\sum_{k\neq i}^{p}\lambda_{ik}\lambda_k - \frac{2}{3}\right)\right]$$

are normalizing transformations of T_{in} in the sense of (2.7). In this case, $g_1(T_{in})$ coincides with the transformation given by Konishi (1981, p.649).

4 Numerical comparisons

We consider the distribution of the i-th largest characteristic root of a sample covariance matrix. Simulated distributions of the normal approximation (N), the approximation based on $g_1(g_1)$, and the approximation based on $g_2(g_2)$ of $n = 30, 50, 100$ are constructed by using $n \times 10^6$ trivariate normal random numbers with population covariance matrix $\mathbf{\Sigma} = \{\sigma_{ij} = 1(i \neq j), \quad \sigma_{ij} = \rho(i = j) : 1 \leq i, j \leq 3\}$. Let $l_{\alpha/2}$ and $l_{1-\alpha/2}$ be the $100(\alpha/2)$ and the $100(1-\alpha/2)$ upper percentiles of the simulated distribution, respectively. Also, let $z_{\alpha/2}$ and $z_{1-\alpha/2}$ be the $100(\alpha/2)$ and the $100(1-\alpha/2)$ upper percentiles of the standard normal distribution, respectively. We investigate the performance of the approximations based on the index

$$Q = |z_{1-\alpha/2} - l_{1-\alpha/2}| + |z_{\alpha/2} - l_{\alpha/2}|.$$

The values of $Q \times 10^5$ for $\rho = 3.0, 5.0, 10.0$ when $\alpha = 0.05$ are listed in Table 1. From Table 1 and the other numerical comparisons, we find that the approximation based on g_1 performs better than the approximation based on g_2 in this example.

5 Concluding Remarks

We proposed two types of concrete normalizing transformations g_1 and g_2. Transformation g_1 is represented as a power function, while g_2 is represented as a exponential function. Our numerical comparisons show that g_1 performs better than g_2 in the example of i-th largest characteristic root of a sample covariance matrix.

Table 1: Values of $(Q \times 10^5)$ when $\alpha = 0.05$

ρ	n	N	g_1	g_2
3.0	30	4718	1452	2023
	50	3705	768	1125
	100	2651	315	505
5.0	30	6043	2720	3177
	50	5098	2090	2395
	100	3845	1033	1204
10.0	30	9660	3729	4096
	50	7774	3368	3631
	100	6038	2726	2874

References

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

[2] Konishi,S. (1991). Normalizing transformations and bootstrap confidence intervals, *Ann. Statist.*, **19**, 2209-2225.

[3] Siotani,M., Hayakawa,T., and Fujikoshi,Y. (1985). *Modern Multivariate Statistical Analysis: A graduate course and handbook. American Science Press.*