Edgeworth Expansions of Functions of the Sample Covariance Matrix with an Unknown Population

(Last Modified: April 24, 2008)

Hirokazu YANAGIHARA[1] AND Ke-Hai YUAN[2]

[1] *Department of Mathematics, Graduate School of Science, Hiroshima University*
1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8626, Japan

[2] *Department of Psychology, University of Notre Dame*
Notre Dame, Indiana 46556, USA

Abstract

By introducing new forms of multivariate cumulants this paper provides the first-order Edgeworth expansions of the standardized and Studentized function of the sample covariance matrix. Without assuming a known population distribution, the obtained expansions are most general and also simpler than those in the literature. A new statistic is also proposed by removing the effect of skewness from that based on standard asymptotics. Because each expansion only involves the first- and second-order derivatives of the function with respect to the sample covariance matrix, the results can be easily applied to many statistics in multivariate analysis. Special cases are also noted when the underlying population follows a normal distribution or an elliptical distribution.

AMS 2000 *subject classifications.* Primary 62H20; Secondary 62F10.
Key words: Asymptotic expansion; Confidence interval, Matrix derivatives, Normal approximation, Removing skewness, Robustness to nonnormality, Testing hypothesis.

1. Introduction

Let $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ be a random sample from a p-variate population $\boldsymbol{y}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The unbiased sample covariance matrix is

$$\boldsymbol{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})', \tag{1}$$

[1] Corresponding author. E-mail: yanagi@math.sci.hiroshima-u.ac.jp

where $\bar{\boldsymbol{y}}$ is the sample mean. Most test statistics in multivariate analysis are functions of $\boldsymbol{S}$. It is important to study the distribution of $h(\boldsymbol{S})$ for a smooth function $h(\cdot)$. The exact distribution of $h(\boldsymbol{S})$ can be obtained for only a few statistics with normally distributed data, e.g., the sample Pearson correlation. These exact distributions are so complicated that they are almost never used in practice. Most inferences for functions of $\boldsymbol{\Sigma}$ are still based on the standard asymptotics, in which the distribution of $h(\boldsymbol{S})$ is approximated by a normal distribution. Such an approximation can be poor when either the sample size is not large enough or $\boldsymbol{y}$ does not follow $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Various methods of improving the normal distribution approximation have been developed (e.g., Sugiura, 1973; Fujikoshi, 1980; Ichikawa & Konishi, 2002; Ogasawara, 2006). These developments either focus on a special statistic or assume $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. In this paper, we will obtain the Edgeworth expansion for $h(\boldsymbol{S})$ and use it to construct a statistic that more closely follows a normal distribution. Because practical data seldom follow a normal distribution, we will not assume any population distribution forms in the development. The development only involves the basic inferential statistics and some algebraic operations. One only needs to identify $h(\boldsymbol{S})$ and calculates the first and second derivatives of $h(\boldsymbol{\Sigma})$ with respect to the elements of $\boldsymbol{\Sigma}$ in order to apply the result to a specific problem.

Section 2 contains some notation and results that will be used for obtaining the asymptotic expansion of the distribution of $h(\boldsymbol{S})$. In Section 3, we give the coefficients in the asymptotic expansions of the distributions of standardized and Studentized $h(\boldsymbol{S})$. In Section 4, we propose a new statistic by removing the effect of skewness from the statistic that is based on standard asymptotics. We also illustrate the application of the new statistic in constructing a better confidence interval for $h(\boldsymbol{\Sigma})$.

2. Preliminary

2.1. Several Higher-order Cumulants

Let

$$\boldsymbol{\varepsilon}_i = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{y}_i - \boldsymbol{\mu}), \ (i = 1, \dots, n).$$

Then $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n$ are independent and identically distributed as $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_p)' = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{y} - \boldsymbol{\mu})$ with $E[\boldsymbol{\varepsilon}] = \boldsymbol{0}$ and covariance matrix $\mathrm{Cov}[\boldsymbol{\varepsilon}] = \boldsymbol{I}_p$. The cumulants of $\boldsymbol{y}$ are often used to obtain the asymptotic expansions of specific functions of $\boldsymbol{S}$. We will use cumulants of $\boldsymbol{\varepsilon}$ for simpler results. In particular, we will introduce several new cumulants of $\boldsymbol{\varepsilon}$ using symmetric matrices as those in Yanagihara (2007).

Let

$$\mu_{a_1\cdots a_j} = E[\varepsilon_{a_1}\cdots\varepsilon_{a_j}]$$

and $\kappa_{a_1\cdots a_j}$ be the corresponding jth-order cumulant of $\boldsymbol{\varepsilon}$. Then there exist

$$\mu_{abc} = \kappa_{abc}, \quad \mu_{abcd} = \kappa_{abcd} + \sum_{[3]} \delta_{ab}\delta_{cd},$$
$$\mu_{abcdef} = \kappa_{abcdef} + \sum_{[10]} \kappa_{abc}\kappa_{def} + \sum_{[15]} \delta_{ab}\kappa_{cdef} + \sum_{[15]} \delta_{ab}\delta_{cd}\delta_{ef},$$

where δ_{ab} is the Kronecker delta, i.e., $\delta_{aa} = 1$ and $\delta_{ab} = 0$ for $a \neq b$; and $\sum_{[j]}$ is the summation of a total of j terms of different combinations, e.g., $\sum_{[3]} \delta_{ab}\delta_{cd} = \delta_{ab}\delta_{cd} + \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}$. Let $\boldsymbol{M} = (m_{ij})$, $\boldsymbol{P} = (p_{ij})$ and $\boldsymbol{Q} = (q_{ij})$ be $p \times p$ symmetric matrices. We define the following multivariate cumulants of the transformed $\boldsymbol{\varepsilon}$ through $\boldsymbol{M}$, $\boldsymbol{P}$ and $\boldsymbol{Q}$:

$$\begin{aligned}
\psi(\boldsymbol{M},\boldsymbol{P}) &= E[(\boldsymbol{\varepsilon}'\boldsymbol{M}\boldsymbol{\varepsilon})(\boldsymbol{\varepsilon}'\boldsymbol{P}\boldsymbol{\varepsilon})] - \{\mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{P}) + 2\mathrm{tr}(\boldsymbol{M}\boldsymbol{P})\} \\
&= \sum_{a,b,c,d}^{p} \kappa_{abcd} m_{ab} p_{cd}, \\
\alpha_1(\boldsymbol{M},\boldsymbol{P},\boldsymbol{Q}) &= E[(\boldsymbol{\varepsilon}_1'\boldsymbol{M}\boldsymbol{\varepsilon}_2)(\boldsymbol{\varepsilon}_1'\boldsymbol{P}\boldsymbol{\varepsilon}_2)(\boldsymbol{\varepsilon}_1'\boldsymbol{Q}\boldsymbol{\varepsilon}_2)] \\
&= \sum_{a,b,c,d,e,f}^{p} \kappa_{abc}\kappa_{def} m_{ad} p_{be} q_{cf}, \\
\alpha_2(\boldsymbol{M},\boldsymbol{P},\boldsymbol{Q}) &= E[(\boldsymbol{\varepsilon}_1'\boldsymbol{M}\boldsymbol{\varepsilon}_1)(\boldsymbol{\varepsilon}_1'\boldsymbol{P}\boldsymbol{\varepsilon}_2)(\boldsymbol{\varepsilon}_2'\boldsymbol{Q}\boldsymbol{\varepsilon}_2)] \\
&= \sum_{a,b,c,d,e,f}^{p} \kappa_{abc}\kappa_{def} m_{ab} p_{cd} q_{ef}, \\
\beta(\boldsymbol{M},\boldsymbol{P},\boldsymbol{Q}) &= E[(\boldsymbol{\varepsilon}'\boldsymbol{M}\boldsymbol{\varepsilon})(\boldsymbol{\varepsilon}'\boldsymbol{P}\boldsymbol{\varepsilon})(\boldsymbol{\varepsilon}'\boldsymbol{Q}\boldsymbol{\varepsilon})] \\
&\quad -2\{2\alpha_1(\boldsymbol{M},\boldsymbol{P},\boldsymbol{Q}) + \alpha_2(\boldsymbol{M},\boldsymbol{P},\boldsymbol{Q}) + \alpha_2(\boldsymbol{M},\boldsymbol{Q},\boldsymbol{P}) + \alpha_2(\boldsymbol{P},\boldsymbol{M},\boldsymbol{Q})\} \\
&\quad -\{\mathrm{tr}(\boldsymbol{M})\psi(\boldsymbol{P},\boldsymbol{Q}) + \mathrm{tr}(\boldsymbol{P})\psi(\boldsymbol{M},\boldsymbol{Q}) + \mathrm{tr}(\boldsymbol{Q})\psi(\boldsymbol{M},\boldsymbol{P})\} \\
&\quad -4\{\psi(\boldsymbol{M},\boldsymbol{P}\boldsymbol{Q}) + \psi(\boldsymbol{P},\boldsymbol{M}\boldsymbol{Q}) + \psi(\boldsymbol{Q},\boldsymbol{M}\boldsymbol{P})\} - \mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{P})\mathrm{tr}(\boldsymbol{Q}) \\
&\quad -2\{\mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{P}\boldsymbol{Q}) + \mathrm{tr}(\boldsymbol{P})\mathrm{tr}(\boldsymbol{M}\boldsymbol{Q}) + \mathrm{tr}(\boldsymbol{Q})\mathrm{tr}(\boldsymbol{M}\boldsymbol{P}) + 4\mathrm{tr}(\boldsymbol{M}\boldsymbol{P}\boldsymbol{Q})\} \\
&= \sum_{a,b,c,d,e,f}^{p} \kappa_{abcdef} m_{ab} p_{cd} q_{ef},
\end{aligned}$$

where the notation $\sum_{a,b,\ldots}^{p}$ means $\sum_{a=1}^{p}\sum_{b=1}^{p}\cdots$. The commonly used multivariate skewnesses and kurtosis (see, Mardia, 1970) are special cases of those defined above, e.g.,

$$\kappa_4^{(1)} = \psi(\boldsymbol{I}_p,\boldsymbol{I}_p), \quad \kappa_{3,3}^{(1)} = \alpha_1(\boldsymbol{I}_p,\boldsymbol{I}_p,\boldsymbol{I}_p), \quad \kappa_{3,3}^{(2)} = \alpha_2(\boldsymbol{I}_p,\boldsymbol{I}_p,\boldsymbol{I}_p).$$

If $\boldsymbol{\varepsilon} \sim N_p(\boldsymbol{0},\boldsymbol{I}_p)$, then all cumulants become 0. If $\boldsymbol{\varepsilon}$ follows an elliptical distribution, there exist

$$\kappa_{abc} = 0, \quad \kappa_{abcd} = \varphi_4 \sum_{[3]} \delta_{ab}\delta_{cd}, \quad \kappa_{abcd} = (\varphi_6 - 3\varphi_4) \sum_{[15]} \delta_{ab}\delta_{cd}\delta_{ef}, \tag{2}$$

where $\varphi_4 = E[\varepsilon_j^4]/3 - 1$ and $\varphi_6 = E[\varepsilon_j^6]/15 - 1$ are the extra kurtosis and 6th-order moments of the jth marginal variate ε_j of $\boldsymbol{\varepsilon}$ relative to those of the standardized normal distribution. Thus, the cumulants of an elliptical distribution are

$$\begin{aligned}
\psi(\boldsymbol{M}, \boldsymbol{P}) &= \varphi_4 \{\mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{P}) + 2\mathrm{tr}(\boldsymbol{MP})\}, \\
\alpha_1(\boldsymbol{M}, \boldsymbol{P}, \boldsymbol{Q}) &= 0, \\
\alpha_2(\boldsymbol{M}, \boldsymbol{P}, \boldsymbol{Q}) &= 0, \\
\beta(\boldsymbol{M}, \boldsymbol{P}, \boldsymbol{Q}) &= (\varphi_6 - 3\varphi_4) \{\mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{P})\mathrm{tr}(\boldsymbol{Q}) + 2\mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{PQ}) \\
&\quad +2\mathrm{tr}(\boldsymbol{P})\mathrm{tr}(\boldsymbol{MQ}) + 2\mathrm{tr}(\boldsymbol{Q})\mathrm{tr}(\boldsymbol{MP}) + 8\mathrm{tr}(\boldsymbol{MPQ})\}.
\end{aligned}$$

For simplicity, we write $\psi(\boldsymbol{M}) = \psi(\boldsymbol{M}, \boldsymbol{M})$, $\alpha_1(\boldsymbol{M}) = \alpha_1(\boldsymbol{M}, \boldsymbol{M}, \boldsymbol{M})$, $\alpha_2(\boldsymbol{M}) = \alpha_2(\boldsymbol{M}, \boldsymbol{M}, \boldsymbol{M})$ and $\beta(\boldsymbol{M}) = \beta(\boldsymbol{M}, \boldsymbol{M}, \boldsymbol{M})$. Then, it follows from the definition that

$$\begin{aligned}
\psi(\boldsymbol{M}) &= E[(\boldsymbol{\varepsilon}'\boldsymbol{M}\boldsymbol{\varepsilon})^2] - \{\mathrm{tr}(\boldsymbol{M})\}^2 - 2\mathrm{tr}(\boldsymbol{M}^2), \\
\alpha_1(\boldsymbol{M}) &= E[(\boldsymbol{\varepsilon}_1'\boldsymbol{M}\boldsymbol{\varepsilon}_2)^2], \\
\alpha_2(\boldsymbol{M}) &= E[(\boldsymbol{\varepsilon}_1'\boldsymbol{M}\boldsymbol{\varepsilon}_1)(\boldsymbol{\varepsilon}_1'\boldsymbol{M}\boldsymbol{\varepsilon}_2)(\boldsymbol{\varepsilon}_2'\boldsymbol{M}\boldsymbol{\varepsilon}_2)], \\
\beta(\boldsymbol{M}) &= E[(\boldsymbol{\varepsilon}'\boldsymbol{M}\boldsymbol{\varepsilon})^3] - 2\{2\alpha_1(\boldsymbol{M}) + 3\alpha_2(\boldsymbol{M})\} - 3\{\mathrm{tr}(\boldsymbol{M})\psi(\boldsymbol{M}) + 4\psi(\boldsymbol{M}, \boldsymbol{M}^2)\} \\
&\quad -\{\mathrm{tr}(\boldsymbol{M})\}^3 - 6\mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{M}^2) - 8\mathrm{tr}(\boldsymbol{M}^3).
\end{aligned}$$

2.2 Standardized and Studentized $h(\boldsymbol{S})$

Let

$$\boldsymbol{V} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i' - \boldsymbol{I}_p), \quad \boldsymbol{z} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\varepsilon}_i.$$

Then both the matrix $\boldsymbol{V}$ and the vector $\boldsymbol{z}$ are asymptotically normally distributed. Using $\boldsymbol{V}$ and $\boldsymbol{z}$, we can expand the $\boldsymbol{S}$ in (1) as

$$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{S}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{I}_p + \frac{1}{\sqrt{n}}\boldsymbol{V} - \frac{1}{n}(\boldsymbol{z}\boldsymbol{z}' - \boldsymbol{I}_p) + O_p(n^{-3/2}). \tag{3}$$

Let

$$\partial_{ij} = \frac{1}{2}(1 + \delta_{ij})\frac{\partial}{\partial \sigma_{ij}},$$

and define $\boldsymbol{\Delta} = (\partial_{ij})$ $(i = 1, \ldots, p; j = 1, \ldots, p)$. Then, we can write the first derivative of $h(\boldsymbol{\Sigma})$ with respect to $\boldsymbol{\Sigma}$ as

$$\boldsymbol{G}(\boldsymbol{M}) = \frac{\partial}{\partial \boldsymbol{\Sigma}} h(\boldsymbol{\Sigma})\bigg|_{\boldsymbol{\Sigma}=\boldsymbol{M}} = \boldsymbol{\Delta} h(\boldsymbol{\Sigma})|_{\boldsymbol{\Sigma}=\boldsymbol{M}} \,. \tag{4}$$

Similarly, letting $\boldsymbol{\delta} = \text{vec}(\boldsymbol{\Delta})$, the first and second derivatives of $h(\boldsymbol{\Sigma})$ with respect to $\boldsymbol{\Sigma}$ are given by

$$\boldsymbol{g}(\boldsymbol{M}) = \boldsymbol{\delta} h(\boldsymbol{\Sigma})|_{\boldsymbol{\Sigma}=\boldsymbol{M}} \,, \qquad \boldsymbol{H}(\boldsymbol{M}) = (\boldsymbol{\delta}\boldsymbol{\delta}')h(\boldsymbol{\Sigma})|_{\boldsymbol{\Sigma}=\boldsymbol{M}} \,. \tag{5}$$

It should be kept in mind that $\boldsymbol{g}(\boldsymbol{M}) = \text{vec}(\boldsymbol{G}(\boldsymbol{M}))$. Let

$$\boldsymbol{v} = \text{vec}(\boldsymbol{V}), \quad \boldsymbol{u} = \text{vec}(\boldsymbol{z}\boldsymbol{z}' - \boldsymbol{I}_p), \quad \boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{1/2} \otimes \boldsymbol{\Sigma}^{1/2}. \tag{6}$$

Applying the Taylor expansion on $h(\boldsymbol{S})$ and using (3) lead to

$$h(\boldsymbol{S}) = h(\boldsymbol{\Sigma}) + \frac{1}{\sqrt{n}} \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Lambda}\boldsymbol{v} + \frac{1}{n}\left\{\frac{1}{2}\boldsymbol{v}'\boldsymbol{\Lambda}\boldsymbol{H}(\boldsymbol{\Sigma})\boldsymbol{\Lambda}\boldsymbol{v} - \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Lambda}\boldsymbol{u}\right\} + O_p(n^{-3/2}). \tag{7}$$

The above expansion will be used to obtain the distribution of $h(\boldsymbol{S})$. We next obtain the standard error of $h(\boldsymbol{S})$.

Let $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{\mu}$ and

$$\boldsymbol{\Omega} = E[\text{vec}(\boldsymbol{r}\boldsymbol{r}' - \boldsymbol{\Sigma})\text{vec}(\boldsymbol{r}\boldsymbol{r}' - \boldsymbol{\Sigma})'].$$

Then $\boldsymbol{\Omega}$ involves the fourth-order cumulants of $\boldsymbol{\varepsilon}$. Let $\boldsymbol{e}_j$ be a $p \times 1$ vector whose jth element is 1 and others are 0, then the $p^2 \times p^2$ matrix

$$\boldsymbol{\Psi} = \sum_{a,b,c,d}^{p} \kappa_{abcd}(\boldsymbol{e}_a\boldsymbol{e}_b' \otimes \boldsymbol{e}_c\boldsymbol{e}_d') \tag{8}$$

contains all the 4th-order cumulants of $\boldsymbol{\varepsilon}$ (Yanagihara, Tonda & Matsumoto, 2005). Let

$$\boldsymbol{K}_p = \sum_{a,b}^{p} (\boldsymbol{e}_a\boldsymbol{e}_b') \otimes (\boldsymbol{e}_b\boldsymbol{e}_a').$$

be the commutation matrix (see Magnus & Neudecker, 1999, p. 48). It follows from

$$\text{vec}(\boldsymbol{r}\boldsymbol{r}' - \boldsymbol{\Sigma}) = \boldsymbol{\Lambda}\text{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)$$

and $\boldsymbol{\Lambda}\boldsymbol{K}_p = \boldsymbol{K}_p\boldsymbol{\Lambda}$ that

$$\boldsymbol{\Omega} = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda} + (\boldsymbol{I}_{p^2} + \boldsymbol{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}). \tag{9}$$

When $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, all the cumulants are zero. Then $\boldsymbol{\Omega}$ becomes $(\boldsymbol{I}_{p^2} + \boldsymbol{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})$. Notice that $\boldsymbol{I}_{p^2} = \sum_{a,b}^{p}(\boldsymbol{e}_a\boldsymbol{e}_a' \otimes \boldsymbol{e}_b\boldsymbol{e}_b')$ and $\text{vec}(\boldsymbol{I}_p)\text{vec}(\boldsymbol{I}_p)' = \sum_{a,b}^{p}(\boldsymbol{e}_a\boldsymbol{e}_b' \otimes \boldsymbol{e}_a\boldsymbol{e}_b')$. When $\boldsymbol{\varepsilon}$ follows an elliptical distribution, $\boldsymbol{\Psi} = \varphi_4\{\boldsymbol{I}_{p^2} + \boldsymbol{K}_p + \text{vec}(\boldsymbol{I}_p)\text{vec}(\boldsymbol{I}_p)'\}$ is obtained by substituting the κ_{abcd} in (2) into (8). This result furher implies that, when $\boldsymbol{\varepsilon}$ follows an elliptical distribution,

$$\boldsymbol{\Omega} = (\varphi_4 + 1)(\boldsymbol{I}_{p^2} + \boldsymbol{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) + \varphi_4 \text{vec}(\boldsymbol{\Sigma})\text{vec}(\boldsymbol{\Sigma})'.$$

It follows from (7) that

$$n\{h(\boldsymbol{S}) - h(\boldsymbol{\Sigma})\}^2 = \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Lambda}\boldsymbol{v}\boldsymbol{v}'\boldsymbol{\Lambda}\boldsymbol{g}(\boldsymbol{\Sigma}) + O_p(n^{-1/2}).$$

Thus,

$$\text{Var}[h(\boldsymbol{S})] = \frac{1}{n}\tau^2 + o(n^{-3/2}),$$

where

$$\tau^2 = \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Omega}\boldsymbol{g}(\boldsymbol{\Sigma}). \tag{10}$$

Since $\boldsymbol{G}(\boldsymbol{\Sigma})$ is symmetric, $\boldsymbol{K}_p\text{vec}(\boldsymbol{G}(\boldsymbol{\Sigma})) = \text{vec}(\boldsymbol{G}(\boldsymbol{\Sigma}))$. Recall that $\boldsymbol{g}(\boldsymbol{\Sigma}) = \text{vec}(\boldsymbol{G}(\boldsymbol{\Sigma}))$. Hence, the τ^2 in (10) can be written as

$$\begin{aligned} \tau^2 &= \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}\boldsymbol{g}(\boldsymbol{\Sigma}) + 2\boldsymbol{g}(\boldsymbol{\Sigma})'(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\boldsymbol{g}(\boldsymbol{\Sigma}) \\ &= \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}\boldsymbol{g}(\boldsymbol{\Sigma}) + 2\text{tr}(\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma})\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma})). \end{aligned}$$

When $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\tau^2 = 2\text{tr}(\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma})\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma}))$. When $\boldsymbol{\varepsilon}$ follows an elliptical distribution, it follows from $\text{vec}(\boldsymbol{\Sigma})'\boldsymbol{g}(\boldsymbol{\Sigma}) = \text{tr}(\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma}))$ that

$$\tau^2 = \varphi_4\{\text{tr}(\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma}))\}^2 + 2(\varphi_4 + 1)\text{tr}(\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma})\boldsymbol{\Sigma}\boldsymbol{G}(\boldsymbol{\Sigma})).$$

Let

$$\hat{\boldsymbol{r}}_i = \boldsymbol{y}_i - \bar{\boldsymbol{y}}. \tag{11}$$

and

$$\hat{\boldsymbol{\Omega}} = \frac{1}{n}\sum_{i=1}^{n} \text{vec}(\hat{\boldsymbol{r}}_i\hat{\boldsymbol{r}}_i' - \boldsymbol{S})\text{vec}(\hat{\boldsymbol{r}}_i\hat{\boldsymbol{r}}_i' - \boldsymbol{S})'. \tag{12}$$

It follows from (10) that

$$\hat{\tau} = \{\boldsymbol{g}(\boldsymbol{S})'\hat{\boldsymbol{\Omega}}\boldsymbol{g}(\boldsymbol{S})\}^{1/2} \tag{13}$$

is consistent for τ. Let

$$T_1 = \frac{\sqrt{n}\{h(\boldsymbol{S}) - h(\boldsymbol{\Sigma})\}}{\tau}, \quad T_2 = \frac{\sqrt{n}\{h(\boldsymbol{S}) - h(\boldsymbol{\Sigma})\}}{\hat{\tau}}. \tag{14}$$

We will call T_1 the standardized $h(\boldsymbol{S})$ and T_2 the Studentized $h(\boldsymbol{S})$. Notice that both T_1 and T_2 are asymptotically distributed according to $N(0,1)$, and there exist

$$P(T_j \leq z_\alpha) = 1 - \alpha + o(1), \quad (j = 1, 2), \tag{15}$$

where $z_\alpha = \Phi^{-1}(1-\alpha)$ with $\Phi(\cdot)$ being the cumulative distribution function of $N(0,1)$. In the next subsection, we will obtain the asymptotic expansions of $P(T_j \le x)$ $(j = 1, 2)$ and use it to improve the normal distribution approximation in (15).

2.3. Edgeworth Expansions of T_1 and T_2

In a typical application, one uses $T_2 \sim N(0,1)$ for inference. But neither T_1 nor T_2 follows $N(0,1)$ exactly. The first and third cumulants of T_j can be expanded as

$$E[T_j] = \frac{1}{\sqrt{n}}\eta_{j,1} + o(n^{-1/2}), \quad E\left[\{T_j - E[T_j]\}^3\right] = \frac{1}{\sqrt{n}}\eta_{j,3} + o(n^{-1/2}). \tag{16}$$

We need the following conditions for the Edgeworth expansions of T_1 and T_2:

- All the 3rd derivatives of $h(\boldsymbol{S})$ are continuous in a neighborhood of $\boldsymbol{S} = \boldsymbol{\Sigma}$, and the 6th-order moments of $\boldsymbol{\varepsilon}$ exist.
- The $p(p+3)/2 \times 1$ vector $\boldsymbol{\xi} = (\boldsymbol{\varepsilon}', \mathrm{vech}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)')'$ satisfies the Cramér's condition

  $$\limsup_{\|\boldsymbol{t}\|\to\infty} |E[\exp(i\boldsymbol{t}'\boldsymbol{\xi})]| < 1,$$

  where $\boldsymbol{t}$ is a $p(p+3)/2 \times 1$ vector and $\|\boldsymbol{t}\|$ is the Euclidean norm of $\boldsymbol{t}$.

It follows from Bhattacharya and Ghosh (1978) and Fujikoshi (1980) that the Edgeworth expansion of T_j is given by

$$P(T_j \le x) = \Phi(x) - \frac{1}{\sqrt{n}}\left\{\eta_{j,1} + \frac{1}{6}\eta_{j,3}(x^2 - 1)\right\}\phi(x) + o(n^{-1/2}), \tag{17}$$

where $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ is the probability density function of $N(0,1)$. Equation (17) implies that the Edgeworth expansion of T_j is determined by its first- and third-order cumulants. We only need to know $\eta_{j,1}$ and $\eta_{j,3}$ to obtain the Edgeworth expansion of T_j.

3. Main Results

3.1. The Standardized $h(\boldsymbol{S})$

We will obtain explicit forms of $\eta_{1,1}$ and $\eta_{1,3}$ in this subsection. For simplicity, we let

$$\boldsymbol{G}_0 = \boldsymbol{\Sigma}^{1/2}\boldsymbol{G}(\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{1/2}, \quad \boldsymbol{g}_0 = \boldsymbol{\Lambda}\boldsymbol{g}(\boldsymbol{\Sigma}), \quad \boldsymbol{H}_0 = \boldsymbol{\Lambda}\boldsymbol{H}(\boldsymbol{\Sigma})\boldsymbol{\Lambda}, \tag{18}$$

where $\boldsymbol{G}(\boldsymbol{\Sigma})$ is given by (4), $\boldsymbol{g}(\boldsymbol{\Sigma})$ and $\boldsymbol{H}(\boldsymbol{\Sigma})$ are given by (5), and $\boldsymbol{\Lambda}$ is given by (6). It follows from (7) and (14) that

$$T_1 = \frac{1}{\tau}\boldsymbol{g}_0'\boldsymbol{v} + \frac{1}{2\tau\sqrt{n}}\left(\boldsymbol{v}'\boldsymbol{H}_0\boldsymbol{v} - 2\boldsymbol{g}_0'\boldsymbol{u}\right) + O_p(n^{-1}),$$

where $\boldsymbol{v}$ and $\boldsymbol{u}$ are given by (6). Let

$$\gamma_1 = \frac{1}{\tau}E[\boldsymbol{v}'\boldsymbol{H}_0\boldsymbol{v}], \qquad \gamma_2 = \frac{\sqrt{n}}{\tau^3}E\left[(\boldsymbol{g}_0'\boldsymbol{v})^3\right],$$
$$\gamma_3 = \frac{1}{\tau^3}E\left[(\boldsymbol{g}_0'\boldsymbol{v})^2\boldsymbol{v}'\boldsymbol{H}_0\boldsymbol{v}\right], \quad \gamma_4 = \frac{1}{\tau^3}E\left[(\boldsymbol{g}_0'\boldsymbol{v})^2\boldsymbol{g}_0'\boldsymbol{u}\right].$$

Then

$$E[T_1] = \frac{1}{2\sqrt{n}}\gamma_1 + o(n^{-1/2}),$$
$$E\left[(T_1 - E[T_1])^3\right] = -\frac{1}{2\sqrt{n}}\left(3\gamma_1 - 2\gamma_2 - 3\gamma_3 + 6\gamma_4\right) + o(n^{-1/2}).$$

Since $\boldsymbol{G}(\boldsymbol{\Sigma})$ is symmetric, $\partial_{ij}\boldsymbol{G}(\boldsymbol{\Sigma})$ is also a symmetric matrix. Notice that

$$\boldsymbol{H}_0 = \boldsymbol{\Lambda}\left(\partial_{11}\boldsymbol{g}(\boldsymbol{\Sigma}), \ldots, \partial_{pp}\boldsymbol{g}(\boldsymbol{\Sigma})\right)\boldsymbol{\Lambda} = \boldsymbol{\Lambda}\left(\mathrm{vec}(\partial_{11}\boldsymbol{G}(\boldsymbol{\Sigma})), \ldots, \mathrm{vec}(\partial_{pp}\boldsymbol{G}(\boldsymbol{\Sigma}))\right)\boldsymbol{\Lambda}.$$

It follows from $\boldsymbol{K}_p\boldsymbol{\Lambda}\mathrm{vec}(\partial_{ij}\boldsymbol{G}(\boldsymbol{\Sigma})) = \boldsymbol{\Lambda}\mathrm{vec}(\partial_{ij}\boldsymbol{G}(\boldsymbol{\Sigma}))$ that $\boldsymbol{K}_p\boldsymbol{H}_0 = \boldsymbol{H}_0$. Also notice that $E[\boldsymbol{v}\boldsymbol{v}'] = \boldsymbol{\Lambda}^{-1}\boldsymbol{\Omega}\boldsymbol{\Lambda}^{-1}$, where $\boldsymbol{\Omega}$ is given by (9). Thus,

$$E[\boldsymbol{v}'\boldsymbol{H}_0\boldsymbol{v}] = \mathrm{tr}(\boldsymbol{\Omega}\boldsymbol{H}(\boldsymbol{\Sigma})) = \mathrm{tr}(\boldsymbol{\Psi}\boldsymbol{H}_0) + 2\mathrm{tr}(\boldsymbol{H}_0),$$

where $\boldsymbol{\Psi}$ is given by (8). Using $\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j - \boldsymbol{I}_p) = \boldsymbol{\varepsilon}_i'\boldsymbol{G}_0\boldsymbol{\varepsilon}_j + \mathrm{tr}(\boldsymbol{G}_0)$, $\boldsymbol{K}_p\boldsymbol{g}_0 = \boldsymbol{g}_0$, and the cumulants introduced in subsection 2.1, we obtain

$$\begin{aligned}
&\sqrt{n}E\left[(\boldsymbol{g}_0'\boldsymbol{v})^3\right] \\
&\quad = E\left[\{\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\}^3\right] \\
&\quad = \beta(\boldsymbol{G}_0) + 4\alpha_1(\boldsymbol{G}_0) + 6\alpha_2(\boldsymbol{G}_0) + 12\psi(\boldsymbol{G}_0, \boldsymbol{G}_0^2) + 8\mathrm{tr}(\boldsymbol{G}_0^3), \\
&E\left[(\boldsymbol{g}_0'\boldsymbol{v})^2\boldsymbol{v}'\boldsymbol{H}_0\boldsymbol{v}\right] \\
&\quad = E\left[\{\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\}^2\right]E\left[\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)'\boldsymbol{H}_0\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\right] \\
&\qquad + 2E\left[\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)'\right]\boldsymbol{H}_0E\left[\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)'\boldsymbol{g}_0\right] + o(1) \\
&\quad = \tau^2\mathrm{tr}(\boldsymbol{\Omega}\boldsymbol{H}(\boldsymbol{\Sigma})) + 2\boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Omega}\boldsymbol{H}(\boldsymbol{\Sigma})\boldsymbol{\Omega}\boldsymbol{g}(\boldsymbol{\Sigma}) + o(1) \\
&\quad = \tau^2\left\{\mathrm{tr}(\boldsymbol{\Psi}\boldsymbol{H}_0) + 2\mathrm{tr}(\boldsymbol{H}_0)\right\} + 2\left(\boldsymbol{g}_0'\boldsymbol{\Psi}\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0\right) + o(1), \\
&E\left[(\boldsymbol{g}_0'\boldsymbol{v})^2\boldsymbol{g}_0'\boldsymbol{u}\right] \\
&\quad = 2E[\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1 - \boldsymbol{I}_p)\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_2 - \boldsymbol{I}_p)\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}_2\boldsymbol{\varepsilon}_2 - \boldsymbol{I}_p)] + o(1) \\
&\quad = 2\alpha_2(\boldsymbol{G}_0) + o(1).
\end{aligned}$$

Combining the above expectations yields

$$\eta_{1,1} = \frac{1}{2\tau}\left\{\mathrm{tr}(\boldsymbol{\Psi}\boldsymbol{H}_0) + 2\mathrm{tr}(\boldsymbol{H}_0)\right\}, \tag{19}$$

$$\begin{aligned}\eta_{1,3} &= \frac{1}{\tau^3}\{3\,(\boldsymbol{g}_0'\boldsymbol{\Psi}\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0') \\ &\quad + \beta(\boldsymbol{G}_0) + 4\alpha_1(\boldsymbol{G}_0) + 12\psi(\boldsymbol{G}_0, \boldsymbol{G}_0^2) + 8\mathrm{tr}(\boldsymbol{G}_0^3)\}.\end{aligned} \tag{20}$$

Let

$$G_j = \mathrm{tr}(\boldsymbol{G}_0^j), \quad \boldsymbol{a} = \mathrm{vec}(\boldsymbol{I}_p). \tag{21}$$

If $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\eta_{1,1}$ and $\eta_{1,3}$ are simplified to

$$\eta_{1,1} = \frac{\mathrm{tr}(\boldsymbol{H}_0)}{(2G_2)^{1/2}}, \quad \eta_{1,3} = \frac{12\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0 + 8G_3}{(2G_2)^{3/2}}.$$

These results coincide with the coefficients in equation (3.5) of Ichikawa and Konishi (2002), who studied the distribution of a standardized $h(\boldsymbol{S})$ under $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. If $\boldsymbol{\varepsilon}$ is distributed according to an elliptical distribution, $\eta_{1,1}$ and $\eta_{1,3}$ are simplified to

$$\begin{aligned}\eta_{1,1} &= \frac{\varphi_4\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{a} + 2(\varphi_4+1)\mathrm{tr}(\boldsymbol{H}_0)}{2\{\varphi_4 G_1^2 + 2(\varphi_4+1)G_2\}^{1/2}}, \\ \eta_{1,3} &= \frac{3\{4(\varphi_4^2+2\varphi_4+1)\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0 + 4(\varphi_4^2+\varphi_4)G_1\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{g}_0 + \varphi_4^2\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{a}\}}{\{\varphi_4 G_1^2 + 2(\varphi_4+1)G_2\}^{3/2}} \\ &\quad + \frac{(\varphi_6 - 3\varphi_4)G_1^3 + 6(\varphi_6-\varphi_4)G_1G_2 + 8(\varphi_6+1)G_3}{\{\varphi_4 G_1^2 + 2(\varphi_4+1)G_2\}^{3/2}}.\end{aligned}$$

3.2. The Studentized $h(\boldsymbol{S})$

This subsection provides explicit forms of $\eta_{2,1}$ and $\eta_{2,3}$. Let

$$\boldsymbol{\Omega}_0 = \boldsymbol{\Lambda}^{-1}\boldsymbol{\Omega}\boldsymbol{\Lambda}^{-1} = \boldsymbol{\Psi} + \boldsymbol{I}_{p^2} + \boldsymbol{K}_p$$

and

$$\boldsymbol{W} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\mathrm{vec}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i' - \boldsymbol{I}_p)\mathrm{vec}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i' - \boldsymbol{I}_p)' - \boldsymbol{\Omega}_0\right\}.$$

Then the matrix $\boldsymbol{W}$ is asymptotically normally distributed. Notice that

$$\boldsymbol{g}(\boldsymbol{S}) = \boldsymbol{g}(\boldsymbol{\Sigma}) + \frac{1}{\sqrt{n}}\boldsymbol{H}(\boldsymbol{\Sigma})\boldsymbol{\Lambda}\boldsymbol{v} + O_p(n^{-1}), \quad \hat{\boldsymbol{\Omega}} = \boldsymbol{\Omega} + \frac{1}{\sqrt{n}}\boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{\Lambda} + O_p(n^{-1}).$$

It follows from (13) and the above expressions that

$$\frac{1}{\hat{\tau}} = \frac{1}{\tau}\left\{1 - \frac{1}{2\tau^2\sqrt{n}}\left(\boldsymbol{g}_0'\boldsymbol{W}\boldsymbol{g}_0 + 2\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\boldsymbol{v}\right)\right\} + O_p(n^{-1}), \tag{22}$$

where $\boldsymbol{g}_0$ and $\boldsymbol{H}_0$ are given by (18). Combining (14) and (22) yields

$$T_2 = T_1 - \frac{1}{2\tau^3\sqrt{n}}\left(\boldsymbol{g}_0'\boldsymbol{W}\boldsymbol{g}_0\boldsymbol{g}_0'\boldsymbol{v} + 2\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\boldsymbol{v}\boldsymbol{g}_0'\boldsymbol{v}\right) + O_p(n^{-1}).$$

Let

$$\gamma_5 = \frac{1}{\tau^3}E[\boldsymbol{g}_0'\boldsymbol{W}\boldsymbol{g}_0\boldsymbol{g}_0'\boldsymbol{v}], \qquad \gamma_6 = \frac{1}{\tau^3}E[\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\boldsymbol{v}\boldsymbol{g}_0'\boldsymbol{v}],$$
$$\gamma_7 = \frac{1}{\tau^5}E[\boldsymbol{g}_0'\boldsymbol{W}\boldsymbol{g}_0(\boldsymbol{g}_0'\boldsymbol{v})^3], \quad \gamma_8 = \frac{1}{\tau^5}E[\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\boldsymbol{v}(\boldsymbol{g}_0'\boldsymbol{v})^3].$$

Then,

$$E[T_2] = \frac{1}{\sqrt{n}}\left(\eta_{1,1} - \frac{1}{2}\gamma_5 - \gamma_6\right) + o(n^{-1/2}), \tag{23}$$

$$E\left[(T_2 - E[T_2])^3\right] = \frac{1}{\sqrt{n}}\left\{\eta_{1,3} + \frac{3}{2}(\gamma_5 - \gamma_7) + 3(\gamma_6 - \gamma_8)\right\} + o(n^{-1/2}), \tag{24}$$

where $\eta_{1,1}$ and $\eta_{1,3}$ are given by (19) and (20), respectively. Using essentially the same technique as for getting the expectations in subsection 3.1, we obtain

$$\begin{aligned}
&E[\boldsymbol{g}_0'\boldsymbol{W}\boldsymbol{g}_0\boldsymbol{g}_0'\boldsymbol{v}]\\
&\quad = E[\{\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\}^3]\\
&\quad = \beta(\boldsymbol{G}_0) + 4\alpha_1(\boldsymbol{G}_0) + 6\alpha_2(\boldsymbol{G}_0) + 12\psi(\boldsymbol{G}_0, \boldsymbol{G}_0^2) + 8\mathrm{tr}(\boldsymbol{G}_0^3),\\
&E[\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\boldsymbol{v}\boldsymbol{g}_0'\boldsymbol{v}]\\
&\quad = \boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Omega}\boldsymbol{H}(\boldsymbol{\Sigma})\boldsymbol{\Omega}\boldsymbol{g}(\boldsymbol{\Sigma})\\
&\quad = \boldsymbol{g}_0'\boldsymbol{\Psi}\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0,\\
&E[\boldsymbol{g}_0'\boldsymbol{W}\boldsymbol{g}_0(\boldsymbol{g}_0'\boldsymbol{v})^3]\\
&\quad = 3E\left[\{\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\}^2\right]E\left[\{\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\}^3\right] + o(1)\\
&\quad = 3\tau^2\left\{\beta(\boldsymbol{G}_0) + 4\alpha_1(\boldsymbol{G}_0) + 6\alpha_2(\boldsymbol{G}_0) + 12\psi(\boldsymbol{G}_0, \boldsymbol{G}_0^2) + 8\mathrm{tr}(\boldsymbol{G}_0^3)\right\} + o(1),\\
&E[\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\boldsymbol{v}(\boldsymbol{g}_0'\boldsymbol{v})^3]\\
&\quad = 3E\left[\{\boldsymbol{g}_0'\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\}^2\right]E\left[\boldsymbol{g}_0'\boldsymbol{\Omega}_0\boldsymbol{H}_0\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)\mathrm{vec}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - \boldsymbol{I}_p)'\boldsymbol{g}_0\right] + o(1)\\
&\quad = 3\tau^2\boldsymbol{g}(\boldsymbol{\Sigma})'\boldsymbol{\Omega}\boldsymbol{H}(\boldsymbol{\Sigma})\boldsymbol{\Omega}\boldsymbol{g}(\boldsymbol{\Sigma}) + o(1)\\
&\quad = 3\tau^2\left(\boldsymbol{g}_0'\boldsymbol{\Psi}\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0\right) + o(1),
\end{aligned}$$

where $\boldsymbol{G}_0$ is given by (18). Using the above expectations in (23) and (24), together with (16) lead to

$$\eta_{2,1} = \frac{1}{2\tau}\left\{\mathrm{tr}(\boldsymbol{\Psi}\boldsymbol{H}_0) + 2\mathrm{tr}(\boldsymbol{H}_0)\right\}$$

$$
\begin{aligned}
& -\frac{1}{2\tau^3}\{2(\boldsymbol{g}_0'\boldsymbol{\Psi}\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0) \\
& \qquad +\beta(\boldsymbol{G}_0) + 4\alpha_1(\boldsymbol{G}_0) + 6\alpha_2(\boldsymbol{G}_0) + 12\psi(\boldsymbol{G}_0, \boldsymbol{G}_0^2) + 8\mathrm{tr}(\boldsymbol{G}_0^3)\}, \\
\eta_{2,3} &= -\frac{1}{\tau^3}\{3(\boldsymbol{g}_0'\boldsymbol{\Psi}\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{\Psi}\boldsymbol{g}_0 + 4\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0') \\
& \qquad +2\beta(\boldsymbol{G}_0) + 8\alpha_1(\boldsymbol{G}_0) + 18\alpha_2(\boldsymbol{G}_0) + 24\psi(\boldsymbol{G}_0, \boldsymbol{G}_0^2) + 16\mathrm{tr}(\boldsymbol{G}_0^3)\}.
\end{aligned}
$$

When $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\eta_{2,1}$ and $\eta_{2,3}$ are simplified to

$$
\eta_{2,1} = \frac{2\{G_2\mathrm{tr}(\boldsymbol{H}_0) - 2\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0\}}{(2G_2)^{3/2}} \quad \text{and} \quad \eta_{2,3} = -\frac{12\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0 + 16G_3}{(2G_2)^{3/2}},
$$

where the G_j's are given by (21). When $\boldsymbol{\varepsilon}$ follows an elliptical distribution, $\eta_{2,1}$ and $\eta_{2,3}$ are simplified to

$$
\begin{aligned}
\eta_{2,1} &= \frac{\varphi_4\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{a} + 2(\varphi_4+1)\mathrm{tr}(\boldsymbol{H}_0)}{2\{\varphi_4G_1^2 + 2(\varphi_4+1)G_2\}^{1/2}} \\
& \quad -\frac{2\{4(\varphi_4^2+2\varphi_4+1)\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0 + 4(\varphi_4^2+\varphi_4)G_1\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{g}_0 + \varphi_4^2\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{a}\}}{2\{\varphi_4G_1^2 + 2(\varphi_4+1)G_2\}^{3/2}} \\
& \quad -\frac{(\varphi_6 - 3\varphi_4)G_1^3 + 6(\varphi_6-\varphi_4)G_1G_2 + 8(\varphi_6+1)G_3}{\{\varphi_4G_1^2 + 2(\varphi_4+1)G_2\}^{3/2}}, \\
\eta_{2,3} &= -\frac{3\{4(\varphi_4^2+2\varphi_4+1)\boldsymbol{g}_0'\boldsymbol{H}_0\boldsymbol{g}_0 + 4(\varphi_4^2+\varphi_4)G_1\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{g}_0 + \varphi_4^2\boldsymbol{a}'\boldsymbol{H}_0\boldsymbol{a}\}}{\{\varphi_4G_1^2 + 2(\varphi_4+1)G_2\}^{3/2}} \\
& \quad -\frac{2\{(\varphi_6 - 3\varphi_4)G_1^3 + 6(\varphi_6-\varphi_4)G_1G_2 + 8(\varphi_6+1)G_3\}}{\{\varphi_4G_1^2 + 2(\varphi_4+1)G_2\}^{3/2}},
\end{aligned}
$$

where $\boldsymbol{a}$ is given by (21).

4. Some Applications

Equation (17) indicates that the approximation $T_2 \sim N(0,1)$ is affected by nonzero $\eta_{2,1}$ and $\eta_{2,3}$. In this section, we propose a new statistic by removing the effect of $\eta_{2,1}$ and $\eta_{2,3}$. Similar statistics in other contexts have been obtained by Hall (1992) and Yanagihara and Yuan (2005).

Let

$$
\begin{aligned}
& c_1 = \mathrm{tr}(\hat{\boldsymbol{\Omega}}\boldsymbol{H}(\boldsymbol{S})), \\
& c_2 = \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{r}}_i'\boldsymbol{G}(\boldsymbol{S})\hat{\boldsymbol{r}}_i - \mathrm{tr}(\boldsymbol{S}\boldsymbol{G}(\boldsymbol{S}))\right\}^3, \\
& c_3 = \boldsymbol{g}(\boldsymbol{S})'\hat{\boldsymbol{\Omega}}\boldsymbol{H}(\boldsymbol{S})\hat{\boldsymbol{\Omega}}\boldsymbol{g}(\boldsymbol{S}), \\
& c_4 = \frac{1}{n^2}\sum_{i,j}^{n}\left\{\hat{\boldsymbol{r}}_i'\boldsymbol{G}(\boldsymbol{S})\hat{\boldsymbol{r}}_i - \mathrm{tr}(\boldsymbol{S}\boldsymbol{G}(\boldsymbol{S}))\right\}\left\{\hat{\boldsymbol{r}}_i'\boldsymbol{G}(\boldsymbol{S})\hat{\boldsymbol{r}}_j - \mathrm{tr}(\boldsymbol{S}\boldsymbol{G}(\boldsymbol{S}))\right\}\left\{\hat{\boldsymbol{r}}_j'\boldsymbol{G}(\boldsymbol{S})\hat{\boldsymbol{r}}_j - \mathrm{tr}(\boldsymbol{S}\boldsymbol{G}(\boldsymbol{S}))\right\},
\end{aligned}
$$

where $\boldsymbol{G}(\boldsymbol{S})$ is given by (4), $\boldsymbol{g}(\boldsymbol{S})$ and $\boldsymbol{H}(\boldsymbol{S})$ are given by (5), and $\hat{\boldsymbol{r}}_i$ and $\hat{\boldsymbol{\Omega}}$ are given by (11) and (12), respectively. Then, consistent estimates of γ_1 to γ_8 are given by

$$\begin{array}{llll} \hat{\gamma}_1 = c_1/\hat{\tau}, & \hat{\gamma}_2 = c_2/\hat{\tau}^3, & \hat{\gamma}_3 = c_1/\hat{\tau} + 2c_3/\hat{\tau}^3, & \hat{\gamma}_4 = 2c_4/\hat{\tau}^3, \\ \hat{\gamma}_5 = c_2/\hat{\tau}^3, & \hat{\gamma}_6 = c_3/\hat{\tau}^3, & \hat{\gamma}_7 = 3c_2/\hat{\tau}^3, & \hat{\gamma}_8 = 3c_3/\hat{\tau}^3, \end{array}$$

where $\hat{\tau}$ is given by (13). It follows from (23) and (24) that

$$\hat{\eta}_{2,1} = \frac{1}{2\hat{\tau}^3}\left(\hat{\tau}^2 c_1 - c_2 - 2c_3\right) \quad \text{and} \quad \hat{\eta}_{2,3} = -\frac{1}{\hat{\tau}^3}(2c_2 + 3c_3 + 6c_4)$$

are consistent for $\eta_{2,1}$ and $\eta_{2,3}$. Let

$$f(x) = x - \frac{1}{6\sqrt{n}}\left\{6\hat{\eta}_{2,1} + \hat{\eta}_{2,3}(x^2 - 1)\right\} + \frac{1}{108n}\hat{\eta}_{2,3}^2 x^3.$$

Then $f(x)$ is monotonically increasing in x. Let

$$T_3 = f(T_2) = T_2 - \frac{1}{6\sqrt{n}}\left\{6\hat{\eta}_{2,1} + \hat{\eta}_{2,3}(T_2^2 - 1)\right\} + \frac{1}{108n}\hat{\eta}_{2,3}^2 T_2^3. \tag{25}$$

It follows from Yanagihara and Yuan (2005) that

$$P(T_3 \le z_\alpha) = 1 - \alpha + o(n^{-1/2}).$$

Thus, using $T_3 \sim N(0,1)$ for inference attains a higher order of accuracy than using $T_2 \sim N(0,1)$.

Many statistical problems in multivariate analysis (see, Tyler, 1983) can be formulated as

$$H_0 : h(\boldsymbol{\Sigma}) = 0 \ \text{ vs } \ H_1 : h(\boldsymbol{\Sigma}) \neq 0.$$

The conventional statistic for testing such a hypothesis is $T_{2,0} = \sqrt{n}h(\boldsymbol{S})/\hat{\tau}$; and, under H_0, there exists

$$P(|T_{2,0}| > z_{\alpha/2}) = \alpha + o(1).$$

Let

$$T_{3,0} = f(T_{2,0}).$$

Then, under H_0,

$$P(|T_{3,0}| > z_{\alpha/2}) = \alpha + o(n^{-1/2}). \tag{26}$$

Thus, $T_{3,0}$ improves the order of accuracy from $o(1)$ in using $T_{2,0}$ to $o(n^{-1/2})$.

The statistic T_3 in (25) also provides a more accurate confidence interval for $h(\boldsymbol{\Sigma})$. The $1-\alpha$ confidence interval for $h(\boldsymbol{\Sigma})$ based on $T_2 \sim N(0,1)$ is given by

$$\mathcal{I}_{1-\alpha}^{(2)} = \left[h(\boldsymbol{S}) - \frac{\hat{\tau}}{\sqrt{n}}z_{\alpha/2},\ h(\boldsymbol{S}) + \frac{\hat{\tau}}{\sqrt{n}}z_{\alpha/2}\right].$$

with

$$P(h(\mathbf{\Sigma}) \in \mathcal{I}_{1-\alpha}^{(2)}) = 1 - \alpha + o(1).$$

When $\hat{\eta}_{2,3} \neq 0$, the inverse of $f(x)$ exists and is given by

$$f^{-1}(x) = \frac{6\sqrt{n}}{\hat{\eta}_{2,3}} + 3\left(\frac{4n}{\hat{\eta}_{2,3}}\left\{x + \frac{1}{6\sqrt{n}}(6\hat{\eta}_{2,1} - \hat{\eta}_{2,3}) - \frac{2\sqrt{n}}{\hat{\eta}_{2,3}}\right\}\right)^{1/3}.$$

The $1 - \alpha$ confidence interval for $h(\mathbf{\Sigma})$ based on $T_3 \sim N(0,1)$ is given by

$$\mathcal{I}_{1-\alpha}^{(3)} = \left[h(\boldsymbol{S}) + \frac{\hat{\tau}}{\sqrt{n}}f^{-1}(-z_{\alpha/2}),\ h(\boldsymbol{S}) + \frac{\hat{\tau}}{\sqrt{n}}f^{-1}(z_{\alpha/2})\right].$$

It follows from (26) and the monotonicity of $f(x)$ that

$$P(h(\mathbf{\Sigma}) \in \mathcal{I}_{1-\alpha}^{(3)}) = 1 - \alpha + o(n^{-1/2}).$$

Thus, the confidence interval using $T_3 \sim N(0,1)$ improve the conventional confidence interval from the order of $o(1)$ with using T_2 to the order of $o(n^{-1/2})$.

References

[1] Bhattacharya, R. N. & Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.*, **6**, 434–451.

[2] Ichikawa, M. & Konishi, S. (2002). Asymptotic expansions and bootstrap approximations in factor analysis. *J. Multivariate Anal.*, **81**, 47–66.

[3] Fujikoshi, Y. (1980). Asymptotic expansions for the distributions of the sample roots under nonnormality. *Biometrika*, **67**, 45–51.

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

[5] Magnus, J. R. & Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics* (revised edition). John Wiley & Sons, New York.

[6] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519–530.

[7] Ogasawara, H. (2006). Asymptotic expansion of the sample correlation coefficient under nonnormality. *Comput. Statist. Data Anal.*, **50**, 891–910.

[8] Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in multivariate analysis. *Comm. Statist.*, **1**, 393–417.

[9] Tyler, D. E. (1983). Robustness and efficiency properties of scatter matrices. *Biometrika*, **70**, 411–20.

[10] Yanagihara, H. (2007). Conditions for robustness to nonnormality on test statistics in a GMANOVA model. *J. Japan Statist. Soc.*, **37**, 135–155,.

[11] Yanagihara, H. & Yuan, K.-H. (2005). Four improved statistics for contrasting means by correcting skewness and kurtosis. *British J. Math. Statist. Psych.*, **58**, 209–237.

[12] Yanagihara, H., Tonda, T. & Matsumoto, C. (2005). The effects of nonnormality on asymptotic distributions of some likelihood ratio criteria for testing covariance structures under normal assumption. *J. Multivariate Anal.*, **96**, No. 2, 237–264.