SUT Journal of Mathematics
Vol. 47, No. 1 (2011), 55–71

A new multivariate kurtosis and its asymptotic distribution

Chiaki Miyagawa, Kazuyuki Koizumi and Takashi Seo

(Received February 27, 2011; Revised June 17, 2011)

Abstract. In this paper, we propose a new definition for multivariate kurtosis based on the two measures of multivariate kurtosis defined by Mardia (1970) and Srivastava (1984), respectively. Under normality, the expectation and the variance for the new multivariate sample measure of kurtosis are given exactly. We also give the third moment for the sample measure of new multivariate kurtosis. After that standardized statistics and normalizing transformation statistic for the sample measure of a new multivariate kurtosis are derived by using these results. Finally, in order to evaluate accuracy of these statistics, we present the numerical results by Monte Carlo simulation for some selected values of parameters.

AMS 2010 *Mathematics Subject Classification.* 62H10, 62D05.

Key words and phrases. Asymptotic distribution, measure of multivariate kurtosis, multivariate normal distribution, normalizing transformation statistic, standardized statistic.

§1. Introduction

In multivariate statistical analysis, the test for multivariate normality is an important problem. This problem has been considered by many authors. Shapiro and Wilk [14] derived test statistic, which is well known as the univariate normality test. This Shapiro-Wilk test was extended for the multivariate case by Malkovich and Afifi [5], Royston [12], Srivastava and Hui [19] and so on. Small [16] gave multivariate extensions of univariate skewness and kurtosis. A comparison of these methods was discussed by Looney [4]. To assess multivariate normality, the sample measures of multivariate skewness and kurtosis have been defined and their null distributions have been given in Mardia [6], [7]. Srivastava [18] also has proposed another definition for the sample measures of multivariate skewness and kurtosis and derived their asymptotic distributions.

Recently, Song [17] has given a definition which is different from Mardia's and Srivastava's measures of multivariate kurtosis. Srivastava's sample measures of multivariate skewness and kurtosis have been discussed by many authors. Seo and Ariga [13] derived the normalizing transformation statistic for Srivastava's sample measure of multivariate kurtosis and its asymptotic distribution. Okamoto and Seo [11] derived the exact values of the expectation and the variance for a sample measure of Srivastava's skewness and improved the approximate χ^2 test statistic for assessing multivariate normality.

On the other hand, Jarque and Bera [2] proposed the bivariate test using skewness and kurtosis for univariate case. The improved Jarque-Bera test statistics have been considered by many authors. For the multivariate case, Mardia and Foster [8] proposed an omnibus test statistic using Mardia's sample measures of skewness and kurtosis. Koizumi, Okamoto and Seo [3] proposed Jarque-Bera type test statistic (MJB) for assessing multivariate normality using Mardia's and Srivastava's skewness and kurtosis. Recently, Enomoto, Okamoto and Seo [1] gave an improved MJB test statistic using Srivastava's skewness and kurtosis.

There are many studies in which the problem for multivariate normality test has been discussed by using skewness and kurtosis. We focus on multivariate kurtosis in this paper. Our purposes are to propose a new definition of multivariate kurtosis from the definition in Mardia [6] and Srivastava [18] and to give the asymptotic distribution. In order to achieve our purposes, we derive the first, second and third moments for a new sample measure of multivariate kurtosis under multivariate normality where the population covariance matrix Σ is known. Further we give the standardized statistics and the normalizing transformation statistic. Finally, we investigate the accuracy of the expectations, the variances, the skewnesses, the kurtosises and the upper percentile for these statistics by Monte Carlo simulation for some selected parameters.

§2. Some definitions of multivariate kurtosis

2.1. Mardia's measure of multivariate kurtosis

First, we discuss a measure of multivariate kurtosis defined by Mardia [6]. Let $\boldsymbol{x}$ be a random p-vector with the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ. Then Mardia [6] has defined the population measure of multivariate kurtosis as

$$\beta_{\mathrm{M}} = \mathrm{E}\left[\left\{(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}^2\right]. \tag{2.1}$$

We note that $\beta_{\mathrm{M}} = p(p+2)$ holds under multivariate normality.

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ be sample observation vectors of size N from a multivariate population. Let $\overline{\boldsymbol{x}} = N^{-1}\sum_{\alpha=1}^{N} \boldsymbol{x}_\alpha$ and $S = N^{-1}\sum_{\alpha=1}^{N}(\boldsymbol{x}_\alpha - \overline{\boldsymbol{x}})(\boldsymbol{x}_\alpha - \overline{\boldsymbol{x}})'$ be the sample mean vector and the sample covariance matrix based on sample size N, respectively. Then the sample measure of multivariate kurtosis in Mardia [6] is defined as

$$b_{\mathrm{M}} = \frac{1}{N}\sum_{\alpha=1}^{N}\left\{(\boldsymbol{x}_\alpha - \overline{\boldsymbol{x}})'S^{-1}(\boldsymbol{x}_\alpha - \overline{\boldsymbol{x}})\right\}^2. \tag{2.2}$$

Further Mardia [6] has obtained asymptotic distributions of b_{M} under the multivariate normality. For the moments and approximation to the null distribution of Mardia's measure of multivariate kurtosis, see, Mardia and Kanazawa [9], Siotani, Hayakawa and Fujikoshi [15].

Theorem 1 (Mardia [6]). *Let b_{M} in (2.2) be the sample measure of multivariate kurtosis on the basis of random samples of size N drawn from $\mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ where Σ is unknown. Then, for large N,*

$$z_{\mathrm{M}} = \frac{(b_{\mathrm{M}} - p(p+2))}{\sqrt{8p(p+2)/N}}$$

is asymptotically distributed as $\mathrm{N}(0,1)$.

2.2. Srivastava's measure of multivariate kurtosis

Next, we consider Srivastava's measure of multivariate kurtosis which is different from the definition by Mardia [6]. Srivastava [18] gave a definition for a measure of kurtosis for multivariate populations using the principle component method. Let $\boldsymbol{x}$ be a random p-vector with the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ. Let $\Gamma = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_p)$ be an orthogonal matrix such that $\Sigma = \Gamma D_\lambda \Gamma'$, where $D_\lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the characteristic roots of Σ.

Then Srivastava [18] defined the population measure of multivariate kurtosis as

$$\beta_{\mathrm{S}} = \frac{1}{p}\sum_{i=1}^{p}\frac{\mathrm{E}[(y_i - \theta_i)^4]}{\lambda_i^2}, \tag{2.3}$$

where $y_i = \boldsymbol{\gamma}_i'\boldsymbol{x}$ and $\theta_i = \boldsymbol{\gamma}_i'\boldsymbol{\mu}$, $i = 1, 2, \ldots, p$. We note that $\beta_{\mathrm{S}} = 3$ holds under multivariate normality. For the moments and approximation to the null distribution of Srivastava's measure of multivariate kurtosis, see, Seo and Ariga [13].

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ be samples of size N from a multivariate population. Let $\overline{\boldsymbol{x}}$ and $S = HD_\omega H'$ be the sample mean vector and the sample covariance matrix based on sample size N, where $H = (\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_p)$ is an orthogonal matrix and $D_\omega = \mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_p)$. We note that $\omega_1, \omega_2, \ldots, \omega_p$ are the characteristic roots of S. Then the sample measure of multivariate kurtosis in Srivastava [18] is defined as

$$b_{\mathrm{S}} = \frac{1}{Np} \sum_{i=1}^{p} \frac{1}{\omega_i^2} \sum_{\alpha=1}^{N} (y_{i\alpha} - \overline{y}_i)^4, \tag{2.4}$$

where $y_{i\alpha} = \boldsymbol{h}_i' \boldsymbol{x}_\alpha$ and $\overline{y}_i = N^{-1} \sum_{\alpha=1}^{N} y_{i\alpha}$, $i = 1, 2, \ldots, p$, $\alpha = 1, 2, \ldots, N$. Further Srivastava (1984) has obtained asymptotic distributions of b_{S} under the multivariate normality.

Theorem 2 (Srivastava [18]). *Let b_{S} in (2.4) be the sample measure of multivariate kurtosis on the basis of random samples of size N drawn from $\mathrm{N}_p(\boldsymbol{\mu}, \ \Sigma)$ where Σ is unknown. Then, for large N,*

$$z_{\mathrm{S}} = \sqrt{\frac{pN}{24}} (b_{\mathrm{s}} - 3)$$

is asymptotically distributed as $\mathrm{N}(0, 1)$.

§3. A new measure of multivariate kurtosis

To propose a new multivariate kurtosis, we reform β_{M} in (2.1) and β_{S} in (2.3) as

$$\beta_{\mathrm{M}} = \mathrm{E}\left[\left\{ (\boldsymbol{x} - \boldsymbol{\mu})' \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right\}^2 \right] = \mathrm{E}\left[\left\{ \mathrm{tr} Z^2 \right\}^2 \right], \tag{3.1}$$

$$\beta_{\mathrm{S}} = \frac{1}{p} \sum_{i=1}^{p} \frac{\mathrm{E}[(y_i - \theta_i)^4]}{\lambda_i^2} = \frac{1}{p} \mathrm{E}\left[\mathrm{tr} Z^4 \right], \tag{3.2}$$

where $Z = \mathrm{diag}(z_1, z_2, \ldots, z_p)$ and $z_i = \lambda_i^{-\frac{1}{2}} (y_i - \theta_i)$, $i = 1, 2, \ldots, p$.

3.1. A new measure of multivariate kurtosis for multivariate populations

Let $\boldsymbol{x}$ be a random p-vector with the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ. From (3.1) and (3.2), we propose that

$$\beta_{\mathrm{MS}} = \frac{1}{p^2} \mathrm{E}\left[\{ \mathrm{tr} Z \}^4 \right].$$

Therefore β_{MS} can be written as

$$\beta_{\mathrm{MS}} = \frac{1}{p^2}\mathrm{E}\left[\{\mathrm{tr}Z\}^4\right] = \frac{1}{p^2}\mathrm{E}\left[\left(\sum_{i=1}^{p}\frac{y_i-\theta_i}{\sqrt{\lambda_i}}\right)^4\right],$$

where, $y_i = \boldsymbol{\gamma}_i'\boldsymbol{x}$ and $\theta_i = \boldsymbol{\gamma}_i'\boldsymbol{\mu}$, $i = 1, 2, \ldots, p$. We note that $\beta_{\mathrm{MS}} = 3$ holds under multivariate normality.

3.2. A sample measure of new multivariate kurtosis

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ be p-dimensional sample vectors of size N from a multivariate population. Then a new sample measure of multivariate kurtosis is defined as

$$b_{\mathrm{MS}} = \frac{1}{Np^2}\sum_{\alpha=1}^{N}\left(\sum_{i=1}^{p}\frac{y_{i\alpha}-\overline{y}_i}{\sqrt{\omega_i}}\right)^4. \tag{3.3}$$

Without loss of generality, we may assume that $\Sigma = I_p$ and $\boldsymbol{\mu} = \boldsymbol{0}$ when we consider this sample measure of multivariate kurtosis. In this paper, we consider the moments for the case when Σ is known under normality. Since we can write $\lambda_i = 1$ $(i = 1, 2, \ldots, p)$ in this case, we can reduce b_{MS} to as follows;

$$b_{\mathrm{MS}} = \frac{1}{Np^2}\sum_{\alpha=1}^{N}\left\{\sum_{i=1}^{p}(y_{i\alpha}-\overline{y}_i)\right\}^4.$$

§4. First moment of b_{MS}

We consider the expectation of b_{MS} under multivariate normality. First we can expand $\mathrm{E}[b_{\mathrm{MS}}]$ given by

$$\begin{aligned}
\mathrm{E}[b_{\mathrm{MS}}] &= \mathrm{E}\left[\frac{1}{Np^2}\sum_{\alpha=1}^{N}\left\{\sum_{i=1}^{p}(y_{i\alpha}-\overline{y}_i)\right\}^4\right] \\
&= \frac{1}{p}\underline{\mathrm{E}[A_{i\alpha}^4]}_{(\mathrm{A})} + \frac{4}{p}(p-1)\underline{\mathrm{E}[A_{i\alpha}^3A_{j\alpha}]}_{(\mathrm{B})} + \frac{3}{p}(p-1)\underline{\mathrm{E}[A_{i\alpha}^2A_{j\alpha}^2]}_{(\mathrm{C})} \\
&\quad + \frac{6}{p}(p-1)(p-2)\underline{\mathrm{E}[A_{i\alpha}^2A_{j\alpha}A_{k\alpha}]}_{(\mathrm{D})} \\
&\quad + \frac{1}{p}(p-1)(p-2)(p-3)\underline{\mathrm{E}[A_{i\alpha}A_{j\alpha}A_{k\alpha}A_{\ell\alpha}]}_{(\mathrm{E})},
\end{aligned} \tag{4.1}$$

where $A_{i\alpha} = y_{i\alpha} - \overline{y}_i$. In order to avoid the dependence of $y_{i\alpha}$ and $\overline{y}_i$, let $\overline{y}_i^{(\alpha)}$ be a mean defined on the subset of $y_{i1}, y_{i2}, \ldots, y_{iN}$ by deleting $y_{i\alpha}$, that is,

$$\overline{y}_i^{(\alpha)} = \frac{1}{N-1} \sum_{j=1, j\neq\alpha}^{N} y_{ij}.$$

Putting $\sqrt{N-1}\,\overline{y}_i^{(\alpha)} = z$, we have

$$A_{i\alpha}^v = \left(1 - \frac{1}{N}\right)^v \left(y_{i\alpha} - \overline{y}_i^{(\alpha)}\right)^v = \left(1 - \frac{1}{N}\right)^v \left(y_{i\alpha} - \frac{z}{\sqrt{N-1}}\right)^v.$$

Then we note that the odd order moments of z and $y_{i\alpha}$ equal zero and

$$\mathrm{E}[z^{2k}] = \mathrm{E}[y_{i\alpha}^{2k}] = (2k-1)\cdots 5 \cdot 3 \cdot 1, \quad k = 1, 2, \ldots.$$

For the case $v = 1, 3, 5, \ldots$, we have $\mathrm{E}\left[A_{i\alpha}^v\right] = 0$. For the case $v = 2k$ ($k = 1, 2, \ldots, 6$), we have

$$\mathrm{E}\left[A_{i\alpha}^{2k}\right] = \left(1 - \frac{1}{N}\right)^k (2k-1)\cdots 5 \cdot 3 \cdot 1.$$

Calculating the cases (A) $, \ldots,$ (E) in (4.1) with respect to $y_{i\alpha}$ and z;

(A) $\mathrm{E}[A_{i\alpha}^4] = 3\left(1 - \frac{1}{N}\right)^2$, (B) $\mathrm{E}[A_{i\alpha}^3 A_{j\alpha}] = 0$,

(C) $\mathrm{E}[A_{i\alpha}^2 A_{j\alpha}^2] = \left(1 - \frac{1}{N}\right)^2$, (D) $\mathrm{E}[A_{i\alpha}^2 A_{j\alpha} A_{k\alpha}] = 0$,

(E) $\mathrm{E}[A_{i\alpha}^2 A_{j\alpha} A_{k\alpha} A_{\ell\alpha}] = 0$,

we obtain

$$\mathrm{E}[b_{\mathrm{MS}}] = 3 - \frac{6}{N} + \frac{3}{N^2}. \tag{4.2}$$

§5. Second and third moments of b_{MS}

In this section, we consider the second and third moments of b_{MS}. We expand $\mathrm{E}\left[b_{\mathrm{MS}}^2\right]$ as follows.

$$
\begin{aligned}
\mathrm{E}\left[b_{\mathrm{MS}}^2\right] &= \frac{1}{N^2p^4}\mathrm{E}\left[\left\{\sum_{\alpha=1}^{N}\left\{\sum_{i=1}^{p}(y_{i\alpha}-\overline{y}_i)\right\}^4\right\}^2\right] \\
&= \frac{1}{N^2p^4}\mathrm{E}\Bigg[\left(\sum_{\alpha=1}^{N}B_\alpha^{(1)}\right)^2+\cdots+\left(\sum_{\alpha=1}^{N}B_\alpha^{(5)}\right)^2 \\
&\quad +2\left(\sum_{\alpha=1}^{N}B_\alpha^{(1)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(2)}\right)+2\left(\sum_{\alpha=1}^{N}B_\alpha^{(1)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(3)}\right) \\
&\quad +2\left(\sum_{\alpha=1}^{N}B_\alpha^{(1)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(4)}\right)+2\left(\sum_{\alpha=1}^{N}B_\alpha^{(1)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(5)}\right) \\
&\quad +2\left(\sum_{\alpha=1}^{N}B_\alpha^{(2)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(3)}\right)+2\left(\sum_{\alpha=1}^{N}B_\alpha^{(2)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(4)}\right) \\
&\quad +2\left(\sum_{\alpha=1}^{N}B_\alpha^{(2)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(5)}\right)+2\left(\sum_{\alpha=1}^{N}B_\alpha^{(3)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(4)}\right) \\
&\quad +2\left(\sum_{\alpha=1}^{N}B_\alpha^{(3)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(5)}\right)+2\left(\sum_{\alpha=1}^{N}B_\alpha^{(4)}\right)\left(\sum_{\alpha=1}^{N}B_\alpha^{(5)}\right)\Bigg],
\end{aligned}
$$

where

$$
\begin{aligned}
B_\alpha^{(1)} &= \sum_{i=1}^{p}A_{i\alpha}^4, \quad B_\alpha^{(2)} = \sum_{i\neq j}^{p}4A_{i\alpha}^3A_{j\alpha}, \quad B_\alpha^{(3)} = \sum_{i<j}^{p}6A_{i\alpha}^2A_{j\alpha}^2, \\
B_\alpha^{(4)} &= \sum_{i,j,k}^{p}12A_{i\alpha}^2A_{j\alpha}A_{k\alpha}, \quad B_\alpha^{(5)} = \sum_{i,j,k,\ell}^{p}A_{i\alpha}A_{j\alpha}A_{k\alpha}A_{\ell\alpha},
\end{aligned}
$$

and $A_{i\alpha} = y_{i\alpha}-\overline{y}_i$. After a great deal of calculation, we obtain

$$
\mathrm{E}\left[b_{\mathrm{MS}}^2\right] = 9+\frac{60}{N}-\frac{258}{N^2}+\frac{324}{N^3}-\frac{135}{N^4}.
$$

Hence we get

$$
\mathrm{Var}\left[b_{\mathrm{MS}}\right] = \frac{96}{N}-\frac{312}{N^2}+\frac{360}{N^3}-\frac{144}{N^4}. \tag{5.1}
$$

For details of these calculations, see Miyagawa and Seo [10]. Next, we consider $\mathrm{E}\left[b_{\mathrm{MS}}^3\right]$ in order to obtain normalizing transformation statistic. As for the normalizing transformation statistic, we discuss in Section 6. Now we can expand $\mathrm{E}[b_{\mathrm{MS}}^3]$ given by

$$\begin{aligned}\mathrm{E}\left[b_{\mathrm{MS}}^3\right] &= \mathrm{E}\left[\left\{\frac{1}{Np^2}\sum_{\alpha=1}^{N}\left(\sum_{i=1}^{p}(y_{i\alpha}-\overline{y}_i)\right)^4\right\}^3\right] \\ &= \frac{1}{N^2p^6}\mathrm{E}\left[\left(\sum_{i=1}^{p}A_{i\alpha}\right)^{12}\right]+\frac{3(N-1)}{N^2p^6}\mathrm{E}\left[\left(\sum_{i=1}^{p}A_{i\alpha}\right)^8\left(\sum_{i=1}^{p}A_{i\beta}\right)^4\right] \\ &\quad+\frac{(N-1)(N-2)}{N^2p^6}\mathrm{E}\left[\left(\sum_{i=1}^{p}A_{i\alpha}\right)^4\left(\sum_{i=1}^{p}A_{i\beta}\right)^4\left(\sum_{i=1}^{p}A_{i\gamma}\right)^4\right],\end{aligned}$$

where $A_{i\alpha}=y_{i\alpha}-\overline{y}_i$. In order to avoid the dependence of $y_{i\alpha},\ y_{i\beta},\ y_{i\gamma}$ and $\overline{y}_i$, let $\overline{y}_i^{(\alpha,\beta,\gamma)}$ be a mean defined on the subset of $y_{i1},y_{i2},\ldots,y_{iN}$ by deleting $y_{i\alpha},\ y_{i\beta}$ and $y_{i\gamma}$, that is,

$$\overline{y}_i^{(\alpha,\beta,\gamma)}=\frac{1}{N-3}\sum_{j=1,\ j\neq\alpha,\beta,\gamma}^{N}y_{ij}.$$

Putting $\sqrt{N-3}\,\overline{y}_i^{(\alpha,\beta,\gamma)}=z$, we have

$$\begin{aligned}&\mathrm{E}\left[(y_{i\alpha}-\overline{y}_i)^u(y_{i\beta}-\overline{y}_i)^v(y_{i\gamma}-\overline{y}_i)^w\right] \\ =&\mathrm{E}\Bigg[\left\{\left(1-\frac{1}{N}\right)y_{i\alpha}-\frac{\sqrt{N-3}}{N}z-\frac{1}{N}y_{i\beta}-\frac{1}{N}y_{\gamma}\right\}^u \\ &\times\left\{\left(1-\frac{1}{N}\right)y_{i\beta}-\frac{\sqrt{N-3}}{N}z-\frac{1}{N}y_{i\gamma}-\frac{1}{N}y_{\alpha}\right\}^v \\ &\times\left\{\left(1-\frac{1}{N}\right)y_{i\gamma}-\frac{\sqrt{N-3}}{N}z-\frac{1}{N}y_{i\alpha}-\frac{1}{N}y_{\beta}\right\}^w\Bigg].\end{aligned}$$

If the values of u,v and w are odd, even and even, respectively, or if all of them are odd, then we have $\mathrm{E}\left[A_{i\alpha}^uA_{i\beta}^vA_{i\gamma}^w\right]=0$. Otherwise, for example, after a great deal of calculation for the expectations, we obtain

$$\begin{aligned}
\mathrm{E}\left[A_{i\alpha}A_{i\beta}A_{i\gamma}^2\right] &= -\frac{1}{N}+\frac{3}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}A_{i\beta}A_{i\gamma}^4\right] &= -\frac{3}{N}+\frac{18}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}^2A_{i\beta}^2A_{i\gamma}^2\right] &= 1-\frac{3}{N}+\frac{9}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}^2A_{i\beta}^2A_{i\gamma}^4\right] &= 3-\frac{12}{N}+\frac{48}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}^2A_{i\beta}^3A_{i\gamma}^3\right] &= -\frac{9}{N}+\frac{45}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}^2A_{i\beta}^4A_{i\gamma}^4\right] &= 9-\frac{45}{N}+\frac{234}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}^3A_{i\beta}^4A_{i\gamma}^4\right] &= 1-\frac{27}{N}+\frac{216}{N^2}+O(N^{-3}),\\
\mathrm{E}\left[A_{i\alpha}^4A_{i\beta}^4A_{i\gamma}^4\right] &= 27-\frac{162}{N}+\frac{1053}{N^2}+O(N^{-3}).
\end{aligned}$$

Therefore the expectation for each term of $\mathrm{E}[b_{\mathrm{MS}}^3]$ is given by

$$\begin{aligned}
\frac{1}{N^2p^6}\mathrm{E}\left[\left(\sum_{i=1}^p A_{i\alpha}\right)^{12}\right] &= \frac{10395}{N^2}+O(N^{-3}),\\
\frac{3(N-1)}{N^2p^6}\mathrm{E}\left[\left(\sum_{i=1}^p A_{i\alpha}\right)^{8}\left(\sum_{i=1}^p A_{i\beta}\right)^{4}\right] &= \frac{945}{N}+\frac{6615}{N^2}+O(N^{-3}),\\
\frac{(N-1)(N-2)}{N^2p^6}\mathrm{E}\left[\left(\sum_{i=1}^p A_{i\alpha}\right)^{4}\left(\sum_{i=1}^p A_{i\beta}\right)^{4}\left(\sum_{i=1}^p A_{i\gamma}\right)^{4}\right]&\\
= 27-\frac{243}{N}+\frac{27(12p^4-72p^3+251p^2-240p+108)}{N^2p^2}+O(N^{-3}).&
\end{aligned}$$

Summarizing these results, we get

$$\begin{aligned}
\mathrm{E}\left[b_{\mathrm{MS}}^3\right] =& 27+\frac{702}{N}\\
&+\frac{27(12p^4-72p^3+391p^2-240p+108)}{N^2p^2}+O(N^{-3}).
\end{aligned} \tag{5.2}$$

§6. Standardized statistics and normalizing transformation statistic for $\mathrm{E}[b_{\mathrm{MS}}]$

By using the results of the expectation and the variance for the sample measures of multivariate kurtosis, we obtain the following theorem.

Theorem 3. *Let* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ *be random samples of size* N *drawn from* $\mathrm{N}_p(\boldsymbol{\mu},\ \Sigma)$*, where* Σ *is known. Then, for large* N*,*

$$z_{\mathrm{MS}} = \sqrt{\frac{N}{96}}(b_{\mathrm{MS}} - 3),$$
$$z^*_{\mathrm{MS}} = \sqrt{\frac{24N^4}{4N^3 - 13N^2 + 15N - 6}} \left\{ b_{\mathrm{MS}} - \left(3 - \frac{6}{N} + \frac{3}{N^2}\right)\right\}$$

are asymptotically distributed as $\mathrm{N}(0,1)$.

Next an asymptotic expansion of the distribution function for a new sample measure of multivariate kurtosis b_{MS} is given under the multivariate normal population. Further, as an improved approximation to a standard normal distribution, we derive the normalizing transformation for the distribution of $\sqrt{N}(b_{\mathrm{MS}} - \beta_{\mathrm{MS}})$.

Let $Y_{\mathrm{MS}} = \sqrt{N}(b_{\mathrm{MS}} - \beta_{\mathrm{MS}})$. Then we have the following distribution function for b_{MS}.

$$\Pr\left[\frac{Y_{\mathrm{MS}}}{\sigma} \le y\right] = \Phi(y) - \frac{1}{\sqrt{N}}\left\{\frac{a_1}{\sigma}\Phi^{(1)}(y) + \frac{a_3}{\sigma^3}\Phi^{(3)}(y)\right\} + O\left(N^{-1}\right),$$

where $\Phi(y)$ is the cumulative distribution function of $\mathrm{N}(0,1)$ and $\Phi^{(j)}(y)$ is the jth derivation of $\Phi(y)$. The first three cumulants of Y_{MS} are given by

$$\kappa_1\left(Y_{\mathrm{MS}}\right) = \frac{a_1}{\sqrt{N}} + O\left(N^{-\frac{3}{2}}\right),$$
$$\kappa_2\left(Y_{\mathrm{MS}}\right) = \sigma^2 + O\left(N^{-1}\right),$$
$$\kappa_3\left(Y_{\mathrm{MS}}\right) = \frac{6a_3}{\sqrt{N}} + O\left(N^{-\frac{3}{2}}\right),$$

where

$$a_1 = 6, \quad \sigma^2 = 96, \quad a_3 = 1584.$$

Further we put the function $g(b_{\mathrm{MS}})$ satisfying the following differential equation

$$\frac{a_3}{\sigma^3} + \frac{\sigma g''(b_{\mathrm{MS}})}{2g'(b_{\mathrm{MS}})} = 0,$$

where $g'(\beta_{\mathrm{S}}) \neq 0$. Solving this equation, we have

$$g(b_{\mathrm{MS}}) = -\frac{32}{11}\exp\left[-\frac{11}{32}b_{\mathrm{MS}}\right].$$

Therefore the above distribution function is transformed as

$$\Pr\left[\frac{\sqrt{N}\left\{g(b_{\rm MS})-g(\beta_{\rm MS}-c/N)\right\}}{\sigma}\le y\right]=\Phi(y)+O\left(N^{-1}\right),$$

where $c=-(45/2)\exp\left[-(11/32)b_{\rm MS}\right]$.

Hence we have following theorem.

Theorem 4. *Let* $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N$ *be random samples of size* N *drawn from* $\mathrm{N}_p(\boldsymbol{\mu},\ \Sigma)$*, where* Σ *is known. Then, for large* N*,*

$$z_{NT}=\frac{\sqrt{N}\left\{-\frac{11}{32}\exp\left[-\frac{11}{32}b_{\rm MS}\right]+\frac{32}{11}\exp\left[-\frac{33}{32}\right]-c/N\right\}}{\sqrt{96}\exp\left[-\frac{33}{32}\right]}$$

is a normalizing transformation for $b_{\rm MS}$*, where* $c=-(45/2)\exp[-(33/32)$ $b_{\rm MS}]$.

§7. Simulation studies

We investigate the accuracy of standardized statistics $z_{\rm MS}$, $z^*_{\rm MS}$ and normalizing transformation statistic z_{NT} by Monte Carlo simulation. Parameters of the dimension and the sample size in simulation are as follows: $p=3,5,7,10,\ N=20,50,100,200,400,800$. As a numerical experiment, we carry out $1,000,000$ replications for the case where $\Sigma(=I_p)$ is known.

Table 1 gives the values of the expectation and the variance for $z_{\rm MS}$, $z^*_{\rm MS}$ and z_{NT}. LT's in Table 1 denote the limiting term for the expectation and the variance of a new multivariate kurtosis. Table 2 gives the values of the skewness and the kurtosis for $z_{\rm MS}$, $z^*_{\rm MS}$ and z_{NT}. LT's in Table 2 denote the limiting term for the skewness and the kurtosis of a new multivariate kurtosis. From Tables 1 and 2, it may be noted that the values for each statistic give good normal approximations when N is large.

It may be seen from Table 1 that the expectation and the variance of all statistics converge to zero and one when N is large. The results show that Theorems 3 and 4 hold. Particularly, the expectations and the variances of the statistic $z^*_{\rm MS}$ are almost same for any N. That is, $z^*_{\rm MS}$ is almost close to the limiting term even for small N, respectively, since $z^*_{\rm MS}$ is a standardized statistic using the exact values of the expectation and the variance derived in this paper. As for the expectation, the accuracy of approximation for z_{NT} is better than that for $z_{\rm MS}$ for any N. Hence, it may be noticed that both $z^*_{\rm MS}$ and z_{NT} are improvement statistics of $z_{\rm MS}$. It may be seen from Table 1 that there is not any effect of dimension at all.

We note that the value of skewness is zero and the value of kurtosis is three under standard normal distribution. The values of skewness and kurtosis for $z_{\rm MS}$ and $z_{\rm MS}^*$ are same since the expectation and the variance are only improved. On the other hand, from Theorem 4, we note that z_{NT} is improved for the distribution function. Therefore it may be noted from Table 2 that the values of skewness and kurtosis for z_{NT} rapidly converge to zero and three. Further it may be seen that the normalizing transformation statistic z_{NT} is pretty good normal approximation even for small N.

Tables 3, 4 and 5 give the upper 10, 5 and 1% points of $z_{\rm MS}$, $z_{\rm MS}^*$ and z_{NT}, respectively. Note that the notation of $z(0.90)$, $z(0.95)$ and $z(0.99)$ mean the upper percent points of normal distribution. In Table 3, $z_{\rm MS}^*$ and z_{NT} are closer to the upper 10% point of normal distribution even when N is small. In Table 4, the accuracy of approximation for $z_{\rm MS}$ is good when N is small. However the upper approximate percent points of $z_{\rm NT}$ are better when N is large. Finally it may be seen from Table 5 that the values for z_{NT} is closer to the upper 1% point of normal distribution for any N.

Some histograms of the sample distributions for $z_{\rm MS}$, $z_{\rm MS}^*$ and z_{NT} by simulation are given in Figure 1 ($p = 10$). Also, we compute the cases $p = 3, 5, 7$, and obtain results similar to these for the case $p = 10$.

In conclusion, it is noted from various points of view that the normalizing transformation statistic improved for the distribution function z_{NT} proposed in this paper is an extremely good normal approximation and is useful for the multivariate normal test.

§8. Conclusion

In this paper, we have proposed a new definition for multivariate kurtosis. It is noticed that the new definition for multivariate kurtosis is based on fourth moment from definitions proposed by Mardia [6] and Srivastava [18]. Under normality, we have derived the expectation, the variance and the third moment for a sample measure of new multivariate kurtosis. Further, standardized statistics and normalizing transformation statistic have been given by using these results. Finally, we have evaluated the accuracy of statistics derived in this paper by Monte Carlo simulation, and we recommend to use z_{NT} for the multivariate normality test. It is left as a future problem for the case when Σ is unknown.

Acknowledgements

The authors would like to thank the referee for his suitable comments and careful reading.

References

[1] Enomoto, R., Okamoto, N. and Seo, T (2010). On the distribution of test statistic using Srivastava's skewness and kurtosis, *Technical Report No.10-07, Statistical Research Group, Hiroshima University.*

[2] Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals, *International Statistical Review*, **55**, 163–172.

[3] Koizumi, K., Okamoto, N. and Seo, T (2009). On Jarque-Bera tests for assessing multivariate normality, *Journal of Statistics: Advances in Theory and Applications*, **1**, 207–220.

[4] Looney, S. W. (1995). How to use test for univariate normality to assess multivariate normality, *American Statistician*, **49**, 64–70.

[5] Malkovich, J. and Afifi, A. A. (1973). On tests for multivariate normality, *Journal of American Statistical Association*, **68**, 176–179.

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

[7] Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, *Sankhyā B*, **36**, 115-128.

[8] Mardia, K. V. and Foster, K. (1983). Omnibus tests of multinormality based on skewness and kurtosis, *Communications in Statistics-Theory and Methods*, **12**, 207–221.

[9] Mardia, K. V. and Kanazawa, M. (1983). The null distribution of multivariate kurtosis, *Communications in Statistics-Simulation and Computation*, **12**, 569–576.

[10] Miyagawa, C. and Seo, T. (2011). A new multivariate kurtosis and its asymptotic distribution, *Technical Report No.11-03, Statistical Research Group, Hiroshima University.*

[11] Okamoto, N. and Seo, T. (2010). On the distribution of multivariate sample skewness, *Journal of Statistical Planning and Inference*, **140**, 2809-2816.

[12] Royston, J. P. (1983). Some techniques for assessing multivariate normality based on the Shapiro-Wilk W, *Applied Statistics*, **32**, 121–133.

[13] Seo, T. and Ariga, M. (2011) On the distribution of kurtosis test for multivariate normality, *Journal of Combinatorics, Information & System Sciences*, **36**, 179–200.

[14] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples), *Biometrika*, **52**, 591–611.

[15] Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*, American Sciences Press, Inc., Columbus, Ohio.

[16] Small, N. J. H. (1980). Marginal skewness and kurtosis in testing multivariate normality, *Applied Statistics*, **22**, 260–266.

[17] Song, K. (2001). Rènyi information, loglikelihood and an intrinsic distribution measure, *Journal of Statistical Planning and Inference*, **93**, 51–69.

[18] Srivastava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality, *Statistics & Probability Letters*, **2**, 263–267.

[19] Srivastava, M. S. and Hui, T. K. (1987). On assessing multivariate normality based on Shapiro-Wilk W statistic, *Statistics & Probability Letters*, **5**, 15–18.

Chiaki Miyagawa
Department of Mathematical Information Science, Graduate School of Science,
Tokyo University of Science, Tokyo, Japan

Kazuyuki Koizumi
Department of Mathematical Information Science, Faculty of Science,
Tokyo University of Science, Tokyo, Japan
E-mail: `zumisj@rs.kagu.tus.ac.jp`

Takashi Seo
Department of Mathematical Information Science, Faculty of Science,
Tokyo University of Science, Tokyo, Japan
E-mail: `seo@rs.kagu.tus.ac.jp`

Table 1: Expectation and variance for $z_{\text{MS}}, z^*_{\text{MS}}$ and z_{NT}

		Expectation (LT:0)			Variance (LT:1)		
	N	z_{MS}	z^*_{MS}	z_{NT}	z_{MS}	z^*_{MS}	z_{NT}
p=3	20	−0.135	−0.001	0.109	0.840	0.991	0.631
	50	−0.086	−0.001	0.036	0.936	1.000	0.794
	100	−0.061	−0.000	0.014	0.970	1.002	0.878
	200	−0.043	−0.001	0.006	0.987	1.003	0.931
	400	−0.029	0.001	0.003	0.992	1.000	0.962
	800	−0.022	0.000	0.000	0.998	1.002	0.983
p=5	20	−0.133	0.000	0.109	0.848	1.002	0.632
	50	−0.085	0.000	0.036	0.938	1.002	0.792
	100	−0.062	−0.001	0.013	0.967	0.999	0.878
	200	−0.044	−0.001	0.005	0.984	1.000	0.931
	400	−0.031	−0.001	0.001	0.990	0.998	0.962
	800	−0.022	0.000	0.000	0.995	0.999	0.979
p=7	20	−0.132	0.002	0.111	0.851	1.005	0.632
	50	−0.085	0.001	0.036	0.939	1.002	0.795
	100	−0.062	−0.001	0.013	0.965	0.997	0.876
	200	−0.044	−0.001	0.004	0.981	0.997	0.929
	400	−0.029	0.001	0.003	0.994	1.002	0.964
	800	−0.022	−0.001	0.000	0.995	0.999	0.980
p=10	20	−1.133	0.000	0.109	0.849	1.003	0.632
	50	−1.085	0.001	0.036	0.940	1.003	0.795
	100	−0.060	0.001	0.014	0.970	1.002	0.877
	200	−0.044	−0.001	0.005	0.983	0.999	0.930
	400	−0.032	0.000	0.001	0.988	0.996	0.960
	800	−0.021	0.000	0.001	0.996	1.000	0.980

Table 2: Skewness and kurtosis for $z_{\text{MS}}, z^*_{\text{MS}}$ and z_{NT}

		Skewness (LT:0)			Kurtosis (LT:3)		
	N	z_{MS}	z^*_{MS}	z_{NT}	z_{MS}	z^*_{MS}	z_{NT}
p=3	20	2.233	2.233	0.068	12.531	12.531	2.193
	50	1.431	1.431	−0.036	7.075	7.075	2.514
	100	1.022	1.022	−0.036	5.113	5.113	2.713
	200	0.724	0.724	−0.014	4.016	4.016	2.845
	400	0.508	0.508	−0.010	3.522	3.522	2.931
	800	0.360	0.360	−0.002	3.254	3.254	2.962
p=5	20	2.276	2.276	0.070	13.262	13.262	2.191
	50	1.456	1.456	−0.030	7.414	7.414	2.518
	100	1.002	1.002	−0.037	4.952	4.952	2.708
	200	0.716	0.716	−0.023	4.025	4.025	2.853
	400	0.502	0.502	−0.014	3.500	3.500	2.927
	800	0.359	0.359	−0.003	3.253	3.253	2.965
p=7	20	2.299	2.299	0.067	13.651	13.651	2.191
	50	1.433	1.433	−0.033	7.024	7.024	2.511
	100	1.006	1.006	−0.037	4.971	4.971	2.713
	200	0.714	0.714	−0.025	4.035	4.035	2.850
	400	0.509	0.509	−0.009	3.514	3.514	2.925
	800	0.353	0.353	−0.010	3.255	3.255	2.970
p=10	20	−1.133	0.000	0.070	13.246	13.246	2.190
	50	−1.085	0.001	−0.032	7.148	7.148	2.512
	100	−0.060	0.001	−0.035	5.043	5.043	2.714
	200	−0.044	−0.001	−0.025	3.999	3.999	2.850
	400	−0.032	0.000	−0.016	3.500	3.500	2.932
	800	−0.021	0.000	−0.007	3.258	3.258	2.968

Table 3: The upper 10% point of $z_{\mathrm{MS}}, z^*_{\mathrm{MS}}$ and z_{NT}

			0.90		
	N	z_{MS}	z^*_{MS}	z_{NT}	$z(0.90)$
p=3	20	0.992	1.223	1.212	1.282
	50	1.160	1.287	1.216	1.282
	100	1.225	1.307	1.233	1.282
	200	1.264	1.318	1.254	1.282
	400	1.277	1.313	1.264	1.282
	800	1.286	1.310	1.273	1.282
p=5	20	0.993	1.225	1.213	1.282
	50	1.163	1.290	1.217	1.282
	100	1.228	1.310	1.235	1.282
	200	1.261	1.315	1.252	1.282
	400	1.275	1.311	1.262	1.282
	800	1.283	1.307	1.271	1.282
p=7	20	0.996	1.228	1.214	1.282
	50	1.161	1.289	1.217	1.282
	100	1.221	1.303	1.231	1.282
	200	1.257	1.310	1.248	1.282
	400	1.282	1.318	1.268	1.282
	800	1.280	1.305	1.269	1.282
p=10	20	0.993	1.224	1.213	1.282
	50	1.162	1.290	1.217	1.282
	100	1.226	1.308	1.234	1.282
	200	1.257	1.311	1.249	1.282
	400	1.269	1.305	1.257	1.282
	800	1.284	1.309	1.272	1.282

Table 4: The upper 5% point of $z_{\mathrm{MS}}, z^*_{\mathrm{MS}}$ and z_{NT}

			0.95		
	N	z_{MS}	z^*_{MS}	z_{NT}	$z(0.95)$
p=3	20	1.605	1.890	1.445	1.645
	50	1.729	1.875	1.503	1.645
	100	1.746	1.837	1.550	1.645
	200	1.751	1.809	1.594	1.645
	400	1.728	1.766	1.614	1.645
	800	1.712	1.737	1.630	1.645
p=5	20	1.608	1.893	1.446	1.645
	50	1.727	1.873	1.502	1.645
	100	1.747	1.838	1.550	1.645
	200	1.744	1.802	1.590	1.645
	400	1.726	1.704	1.613	1.645
	800	1.712	1.737	1.630	1.645
p=7	20	1.614	1.899	1.448	1.645
	50	1.727	1.874	1.502	1.645
	100	1.745	1.835	1.549	1.645
	200	1.738	1.796	1.586	1.645
	400	1.730	1.768	1.616	1.645
	800	1.705	1.730	1.625	1.645
p=10	20	1.612	1.897	1.447	1.645
	50	1.731	1.878	1.504	1.645
	100	1.750	1.841	1.552	1.645
	200	1.741	1.799	1.588	1.645
	400	1.723	1.760	1.610	1.645
	800	1.708	1.733	1.627	1.645

Table 5: The upper 1% point of $z_{\mathrm{MS}}, z^*_{\mathrm{MS}}$ and z_{NT}

			0.99		
	N	z_{MS}	z^*_{MS}	z_{NT}	$z(0.99)$
p=3	20	3.161	3.581	1.719	2.326
	50	3.066	3.257	1.937	2.326
	100	2.925	3.035	2.090	2.326
	200	2.797	2.864	2.204	2.326
	400	2.664	2.706	2.262	2.326
	800	2.574	2.601	2.298	2.326
p=5	20	3.202	3.625	1.722	2.326
	50	3.074	3.265	1.939	2.326
	100	2.917	3.025	2.087	2.326
	200	2.773	2.848	2.197	2.326
	400	2.665	2.693	2.253	2.326
	800	2.567	2.594	2.293	2.326
p=7	20	3.168	3.610	1.721	2.326
	50	3.079	3.270	1.940	2.326
	100	2.917	3.027	2.087	2.326
	200	2.773	2.839	2.192	2.326
	400	2.665	2.706	2.262	2.326
	800	2.559	2.596	2.287	2.326
p=10	20	3.203	3.626	1.722	2.326
	50	3.076	3.267	1.939	2.326
	100	2.939	3.049	2.095	2.326
	200	2.772	2.838	2.192	2.326
	400	2.654	2.695	2.255	2.326
	800	2.561	2.588	2.289	2.326