

ESTIMATION OF MOMENT PARAMETER IN ELLIPTICAL DISTRIBUTIONS

Yosihito Maruyama* and Takashi Seo**

As a typical non-normal case, we consider a family of elliptically symmetric distributions. Then, the moment parameter and its consistent estimator are presented. Also, the asymptotic expectation and the asymptotic variance of the consistent estimator of the general moment parameter are given. Besides, the numerical results obtained by Monte Carlo simulation for some selected parameters are provided.

Key words and phrases: Asymptotic expansion, consistent estimator, elliptical distribution, kurtosis parameter, moment parameter, Monte Carlo simulation, perturbation method.

1. Introduction

The general moment parameter includes the important kurtosis parameter in the study of multivariate statistical analysis for elliptical populations. The kurtosis parameter, especially with relation to the estimation problem, has been considered by many authors. Mardia (1970, 1974) defined a measure of multivariate sample kurtosis and derived its asymptotic distribution for samples from a multivariate normal population. Also, the testing normality was considered by using the asymptotic result. The related discussion of the kurtosis parameter under the elliptical distribution has been given by Anderson (1993), and Seo and Toyama (1996). Henze (1994) has discussed the asymptotic variance of the multivariate sample kurtosis for general distributions. Here we deal with the estimation of the general moment parameters in elliptical distributions. In particular, we make a generalization of the results of Anderson (1993) and give an extension of asymptotic properties in Mardia (1970, 1974), Seo and Toyama (1996). In general, it is not easy to derive the exact distribution of test statistics or the percentiles for the testing problem under the elliptical populations, and so the asymptotic expansion of the statistics is considered. Especially that given up to the higher order includes not only the kurtosis parameter but the more general higher order moment parameters as well. Then we have to speculate about the estimation of the moment parameters as a practical problem. The present paper is organized in the following way. First, the probability density function, characteristic function and subclasses of the elliptical distribution are explained. Secondly, we prove the results concerning the moments and define the moment

Received October 28, 2002. Revised April 21, 2003. Accepted October 7, 2003.

*Department of Mathematics, Graduate School of Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan.

**Department of Mathematical Information Sciences, Faculty of Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan.

parameter. These are done in Section 2. Although Berkane and Bentler (1986, 1987) computed the moments of an elliptically distributed random vector in a different way, we can confirm that our results coincide with theirs. In Section 3, we construct the consistent estimator of the general moment parameters by the moment method. In addition, the asymptotic properties of those estimators for the elliptical distribution are presented in the two cases for which the population covariance matrix is known and unknown. From the results, another estimator that has bias correction with respect to the sample size is proposed. In the last instance, we investigate the bias and the MSE of the estimators by Monte Carlo simulation applied to some selected parameters, and moreover evaluate the confidence intervals for the general moment parameter with the asymptotic properties in Section 4.

2. Moments for elliptical population

Let the p -variate random vector \mathbf{X} be distributed as the p -variate elliptical distribution with parameters $\boldsymbol{\mu}$ and Λ , which is denoted by $E_p(\boldsymbol{\mu}, \Lambda)$, where Λ is some positive definite symmetric matrix. If the probability density function of \mathbf{X} exists, it has the form

$$f(\mathbf{x}) = c_p |\Lambda|^{-1/2} g((\mathbf{x} - \boldsymbol{\mu})' \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu}))$$

for some non-negative function g , where c_p is positive constant. The characteristic function is of the form

$$\phi(\boldsymbol{\theta}) = \exp[i\boldsymbol{\theta}'\boldsymbol{\mu}] \psi(\boldsymbol{\theta}'\Lambda\boldsymbol{\theta})$$

for some function ψ , where $i = \sqrt{-1}$. Note that the expectation and the covariance matrix of \mathbf{X} are $E(\mathbf{X}) = \boldsymbol{\mu}$ and $Cov(\mathbf{X}) = \Sigma = -2\psi'(0)\Lambda$, respectively.

Elliptical distributions include the several special cases, for instance, the multivariate normal distribution which is described as $N_p(\boldsymbol{\mu}, \Lambda)$ with the probability density function

$$f(\mathbf{x}) = (2\pi)^{-p/2} |\Lambda|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu})\right],$$

the multivariate t distribution with ν degrees of freedom and the probability density function

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) (\nu\pi)^{p/2}} |\Lambda|^{-1/2} \left(1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})' \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)^{-(\nu+p)/2},$$

the contaminated normal distribution with the probability density function

$$f(\mathbf{x}) = \frac{1-\omega}{(2\pi)^{p/2}} |\Lambda|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu})\right] \\ + \frac{\omega}{(2\pi\tau^2)^{p/2}} |\Lambda|^{-1/2} \exp\left[-\frac{1}{2\tau^2}(\mathbf{x} - \boldsymbol{\mu})' \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu})\right], \quad 0 \leq \omega \leq 1,$$

and so on, which are referred to e.g., Cambanis *et al.* (1981), Muirhead (1982).

Now we will calculate the moments of elliptically distributed random variates. Suppose that A is a nonsingular matrix satisfying $A'A = \Lambda$. Then we can express $A^{-1}(\mathbf{X} - \boldsymbol{\mu}) = R\mathbf{U}$, where the vector \mathbf{U} has the uniform distribution on the unit sphere $\mathbf{U}'\mathbf{U} = 1$ such that \mathbf{U} is independent of the scalar R which is non-negative. The central moments of \mathbf{X} can be found from the moments of R and \mathbf{U} . In normal case, R^2 has the chi-square distribution with p degrees of freedom. So the m -th order moments of R^2 are given as

$$E(R^{2m}) = 2^m \left(\frac{p}{2}\right)_m,$$

where

$$\left(\frac{p}{2}\right)_m \equiv \left(\frac{p}{2}\right) \left(\frac{p}{2} + 1\right) \cdots \left(\frac{p}{2} + m - 1\right).$$

Then, for example, we have the second, fourth and sixth order moments of \mathbf{U} given by

$$\begin{aligned} E(\mathbf{U}\mathbf{U}') &= \frac{1}{p}I_p, \\ E(U_iU_jU_kU_l) &= \frac{1}{p(p+2)} \sum_{(3)} \delta_{ij}\delta_{kl}, \\ E(U_iU_jU_kU_lU_sU_t) &= \frac{1}{p(p+2)(p+4)} \sum_{(15)} \delta_{ij}\delta_{kl}\delta_{st}, \end{aligned}$$

where δ_{ij} is the Kronecker's delta, and the sum occurs over all ways of grouping the subscripts,

$$\begin{aligned} \sum_{(3)} \delta_{ij}\delta_{kl} &= \delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}, \\ \sum_{(15)} \delta_{ij}\delta_{kl}\delta_{st} &= \delta_{ij}\delta_{kl}\delta_{st} + \delta_{ij}\delta_{ks}\delta_{lt} + \delta_{ij}\delta_{kt}\delta_{ls} + \delta_{ik}\delta_{jl}\delta_{st} + \delta_{ik}\delta_{js}\delta_{lt} \\ &\quad + \delta_{ik}\delta_{jt}\delta_{ls} + \delta_{il}\delta_{jk}\delta_{st} + \delta_{il}\delta_{js}\delta_{kt} + \delta_{il}\delta_{jt}\delta_{ks} + \delta_{is}\delta_{jk}\delta_{lt} \\ &\quad + \delta_{is}\delta_{jl}\delta_{kt} + \delta_{is}\delta_{jt}\delta_{kl} + \delta_{it}\delta_{jk}\delta_{ls} + \delta_{it}\delta_{jl}\delta_{ks} + \delta_{it}\delta_{js}\delta_{kl}. \end{aligned}$$

In general, we have the following result.

LEMMA 1. *The odd order moments of \mathbf{U} are 0, the $2m$ -th order moments are of the form*

$$E(U_iU_jU_kU_l \cdots U_uU_v) = \frac{1}{2^m \left(\frac{p}{2}\right)_m} \sum_{(d_m)} \delta_{ij}\delta_{kl} \cdots \delta_{uv},$$

where

$$d_m = 2^m \left(\frac{1}{2}\right)_m.$$

On the other hand, the $2m$ -th order moments of R in the elliptical distributions are given by

$$E(R^{2m}) = (-4)^m \left(\frac{\rho}{2}\right)_m \psi^{(m)}(0),$$

which is due to Hayakawa and Puri (1985). From them, we have the following result.

LEMMA 2. *For elliptical population, the odd order moments of $\mathbf{X} - \boldsymbol{\mu}$ are 0, the $2m$ -th order moments are of the form*

$$E\left[(X_i - \mu_i)(X_j - \mu_j) \cdots (X_u - \mu_u)(X_v - \mu_v)\right] = (\mathcal{K}_{(m)} + 1) \sum_{(d_m)} \sigma_{ij} \sigma_{kl} \cdots \sigma_{uv},$$

where $\Sigma = (\sigma_{ij})$, and

$$\mathcal{K}_{(m)} \equiv \frac{\psi^{(m)}(0)}{\{\psi'(0)\}^m} - 1, \quad d_m = 2^m \left(\frac{1}{2}\right)_m.$$

The result of Lemma 2 coincides with which Berkane and Bentler (1987) derived by successive differentiations of $\phi(\boldsymbol{\theta})$, and also when $i = j = \cdots = u = v$, we have the moments which are the same as that given by Berkane and Bentler (1986). Here we define $\mathcal{K}_{(m)}$ as the $2m$ -th order moment parameter. Especially, $\mathcal{K}_{(2)}$ is simply denoted by κ and called a kurtosis parameter. The elliptical distributions are characterized by a kurtosis parameter which is included in the general moment parameter.

From the relation between the moments and the cumulants (e.g., see Kendall *et al.* (1987)), for example, the fourth, sixth and eighth cumulants are given by

$$\begin{aligned} \kappa_{ijkl} &= \kappa \sum_{(3)} \sigma_{ij} \sigma_{kl}, \\ \kappa_{ijklst} &= (\mathcal{K}_{(3)} - 3\kappa) \sum_{(15)} \sigma_{ij} \sigma_{kl} \sigma_{st}, \\ \kappa_{ijklstuv} &= (\mathcal{K}_{(4)} - 4\mathcal{K}_{(3)} - 3\kappa^2 + 6\kappa) \sum_{(105)} \sigma_{ij} \sigma_{kl} \sigma_{st} \sigma_{uv}. \end{aligned}$$

3. Asymptotic properties

In this section, we give careful consideration to the consistent estimators of the general moment parameter in the elliptical distributions by the moment method. We also give the asymptotic expectation and the asymptotic variance of those estimators.

Suppose that $\mathbf{X}_1, \dots, \mathbf{X}_n$ are independent and identically distributed random vectors according to $E_p(\boldsymbol{\mu}, \Lambda)$. A multivariate coefficient of kurtosis in the

sense of Mardia (1970) is $\beta_{2,p} \equiv E [\{(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})\}^2]$, and the affine invariant sample analogue of $\beta_{2,p}$ is obtained by

$$\hat{\beta}_{2,p} \equiv \frac{1}{n} \sum_{i=1}^n \{(\mathbf{X}_i - \bar{\mathbf{X}})' S^{-1} (\mathbf{X}_i - \bar{\mathbf{X}})\}^2,$$

where $\bar{\mathbf{X}}$ is the sample mean vector. S is the sample covariance matrix,

$$S \equiv \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$$

Now $\beta_{2,p}$ can be calculated as $p(p+2)(\kappa+1)$ with Lemma 2. For the consistent estimator of κ , we can propose

$$\hat{\kappa} = \frac{1}{p(p+2)} \hat{\beta}_{2,p} - 1.$$

Further, as generality of this result,

$$\beta_{m,p} \equiv E [\{(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})\}^m] = 2^m \left(\frac{p}{2}\right)_m (\mathcal{K}_{(m)} + 1).$$

Therefore, we have a consistent estimator of the $2m$ -th order moment parameter given by

$$(3.1) \quad \hat{\mathcal{K}}_{(m)} = \frac{1}{2^m \left(\frac{p}{2}\right)_m} \hat{\beta}_{m,p} - 1,$$

where

$$\hat{\beta}_{m,p} = \frac{1}{n} \sum_{i=1}^n \{(\mathbf{X}_i - \bar{\mathbf{X}})' S^{-1} (\mathbf{X}_i - \bar{\mathbf{X}})\}^m.$$

This representation is an extension of the estimator of kurtosis parameter discussed by Anderson (1993). We first consider the asymptotic properties of the consistent estimator, which is described by $\hat{\mathcal{K}}_{(m)}^*$ (or $\hat{\beta}_{m,p}^*$) when Σ is known. In this case, it is assumed without loss of generality that $\Sigma = I_p$. Then replacing S with Σ , we can write

$$\hat{\beta}_{m,p}^* = \frac{1}{n} \sum_{i=1}^n T_i^{2m},$$

where $T_i^2 = (\mathbf{X}_i - \bar{\mathbf{X}})' (\mathbf{X}_i - \bar{\mathbf{X}})$. In order to avoid the dependence of \mathbf{X}_i and $\bar{\mathbf{X}}$, we define

$$\bar{\mathbf{X}}_{(i)} = \frac{1}{n-1} \sum_{k \neq i}^n \mathbf{X}_k.$$

Then, we can write

$$T_i^2 = \left(1 - \frac{1}{n}\right)^2 (\mathbf{X}_i - \overline{\mathbf{X}}_{(i)})' (\mathbf{X}_i - \overline{\mathbf{X}}_{(i)}).$$

Note that \mathbf{X}_i is independent of $\overline{\mathbf{X}}_{(i)}$. To obtain the expectation of $\hat{\mathcal{K}}_{(m)}$ by the perturbation method, we put

$$\overline{\mathbf{X}}_{(i)} = \frac{1}{\sqrt{n-1}} \mathbf{Y},$$

then, T_i^{2m} can be expanded as

$$T_i^{2m} = (\mathbf{X}'_i \mathbf{X}_i)^m + \frac{1}{\sqrt{n}} \gamma_1 + \frac{1}{n} \gamma_2 + O_p(n^{-3/2}),$$

where

$$\gamma_1 = -2m (\mathbf{X}'_i \mathbf{X}_i)^{m-1} \mathbf{Y}' \mathbf{X}_i,$$

$$\gamma_2 = m (\mathbf{X}'_i \mathbf{X}_i)^{m-1} (\mathbf{Y}' \mathbf{Y} - 2\mathbf{X}'_i \mathbf{X}_i) + 2m(m-1) (\mathbf{X}'_i \mathbf{X}_i)^{m-2} (\mathbf{Y}' \mathbf{X}_i)^2.$$

Therefore, calculating the expectation with respect to \mathbf{X}_i and \mathbf{Y} , we get the asymptotic mean of $T = \sqrt{n}(\hat{\mathcal{K}}_{(m)}^* - \mathcal{K}_{(m)})$ up to the order n^{-1} as

$$(3.2) \quad E(T) = \frac{1}{\sqrt{n}} \left\{ -m (2\mathcal{K}_{(m)} - \mathcal{K}_{(m-1)} + 1) \right\} + O(n^{-3/2}).$$

If the underlying distribution is normal then $\mathcal{K}_{(m)} = 0$, the asymptotic expectation of T can be reduced as

$$E(T) = -\frac{m}{\sqrt{n}} + O(n^{-3/2}).$$

Similarly, calculating the expectation for the expansion of T_i^{4m} , we may come by the variance. Therefore, we have the following result.

THEOREM 1. *Suppose that $\hat{\mathcal{K}}_{(m)}^*$ is the consistent estimator of the $2m$ -th order moment parameter $\mathcal{K}_{(m)}$ with known Σ , which is given by (3.1). Then the asymptotic variance of $T = \sqrt{n}(\hat{\mathcal{K}}_{(m)}^* - \mathcal{K}_{(m)})$ is*

$$(3.3) \quad \sigma_T^2 = \frac{\binom{\frac{p}{2} + m}{2}_m}{\binom{\frac{p}{2}}{2}_m} (\mathcal{K}_{(2m)} + 1) - (\mathcal{K}_{(m)} + 1)^2.$$

Here, we correct the bias of the estimator $\hat{\mathcal{K}}_{(m)}^*$ as follows. The modified estimator can be obtained by

$$(3.4) \quad \tilde{\mathcal{K}}_{(m)}^* = \hat{\mathcal{K}}_{(m)}^* + \frac{m}{n} \left(2\hat{\mathcal{K}}_{(m)}^* - \hat{\mathcal{K}}_{(m-1)}^* + 1 \right).$$

Note that the bias in $\tilde{\mathcal{K}}_{(m)}^*$ is order n^{-2} with the same variance up to the order n^{-1} as that of $\hat{\mathcal{K}}_{(m)}^*$.

Next, we consider the case when Σ is unknown, that is the main purpose of this paper. In order to avoid the dependence of \mathbf{X}_i , $\bar{\mathbf{X}}$ and S , we put

$$\begin{aligned} T_i^2 &= (\mathbf{X}_i - \bar{\mathbf{X}})' S^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}), \\ \bar{\mathbf{X}}_{(i)} &= \frac{1}{n-1} \sum_{k \neq i}^n \mathbf{X}_k, \\ S_{(i)} &= \frac{1}{n-1} \sum_{k \neq i}^n (\mathbf{X}_k - \bar{\mathbf{X}}_{(i)}) (\mathbf{X}_k - \bar{\mathbf{X}}_{(i)})'. \end{aligned}$$

Then, we can write

$$S = \frac{n-1}{n} \left\{ S_{(i)} + \frac{1}{n} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)}) (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)})' \right\},$$

and also

$$S^{-1} = \frac{n}{n-1} S_{(i)}^{-1} - \frac{\frac{1}{n-1} S_{(i)}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)}) (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)})' S_{(i)}^{-1}}{1 + \frac{1}{n} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)})' S_{(i)}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)})}.$$

Therefore, we note that

$$T_i^2 = \left(1 - \frac{1}{n}\right)^2 \frac{\tilde{T}_i^2}{1 + \frac{n-1}{n^2} \tilde{T}_i^2},$$

where

$$\tilde{T}_i^2 = \frac{n}{n-1} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)})' S_{(i)}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i)}).$$

Further, let

$$\begin{aligned} \bar{\mathbf{X}}_{(i)} &= \frac{1}{\sqrt{n-1}} \mathbf{Y}, \\ S_{(i)} &= I_p + \frac{1}{\sqrt{n-1}} M, \end{aligned}$$

then, we can expand \tilde{T}_i^2 as

$$\tilde{T}_i^2 = \mathbf{X}_i' \mathbf{X}_i + \frac{1}{\sqrt{n}} a_1 + \frac{1}{n} a_2 + O_p(n^{-3/2}),$$

where

$$\begin{aligned} a_1 &= -2\mathbf{X}'_i\mathbf{Y} - \mathbf{X}'_iM\mathbf{X}_i, \\ a_2 &= 2\mathbf{X}'_i\mathbf{X}_i + 2\mathbf{X}'_iM\mathbf{Y} + \mathbf{X}'_iM^2\mathbf{X}_i + \mathbf{Y}'\mathbf{Y}. \end{aligned}$$

Therefore, T_i^{2m} is expanded as

$$T_i^{2m} = (\mathbf{X}'_i\mathbf{X}_i)^m + \frac{1}{\sqrt{n}}b_1 + \frac{1}{n}b_2 + O_p\left(n^{-3/2}\right),$$

where

$$\begin{aligned} b_1 &= -m(\mathbf{X}'_i\mathbf{X}_i)^{m-1}a_1, \\ b_2 &= \frac{m}{2}(\mathbf{X}'_i\mathbf{X}_i)^{m-2}\left\{(m-1)a_1^2 + 2\mathbf{X}'_i\mathbf{X}_ia_2 - 4(\mathbf{X}'_i\mathbf{X}_i)^2 - 2(\mathbf{X}'_i\mathbf{X}_i)^3\right\}. \end{aligned}$$

By using the asymptotic expanded joint probability density function of \mathbf{Y} and M (e.g., see Wakaki (1997)), we can calculate the expectation for the expansion of T_i^{2m} . Then we obtain the asymptotic mean of $T = \sqrt{n}(\hat{\mathcal{K}}_{(m)} - \mathcal{K}_{(m)})$ up to the order n^{-1} as

$$(3.5) \quad E(T) = \frac{1}{\sqrt{n}}c + O\left(n^{-3/2}\right),$$

where

$$\begin{aligned} c &= m(\mathcal{K}_{(m-1)} + 1) - m(p + 2m)(\mathcal{K}_{(m+1)} + 1) \\ &\quad + (3 + mp + 2m)(\kappa + 1)(\mathcal{K}_{(m)} + 1) - (m + 1)(\mathcal{K}_{(m)} + 1). \end{aligned}$$

Similarly, we may come by the variance. In order to avoid the dependence of \mathbf{X}_i , \mathbf{X}_j , $\bar{\mathbf{X}}$ and S , we define

$$\begin{aligned} \bar{\mathbf{X}}_{(i,j)} &= \frac{1}{n-2} \sum_{k \neq i,j}^n \mathbf{X}_k, \\ S_{(i,j)} &= \frac{1}{n-2} \sum_{k \neq i,j}^n (\mathbf{X}_k - \bar{\mathbf{X}}_{(i,j)})(\mathbf{X}_k - \bar{\mathbf{X}}_{(i,j)})'. \end{aligned}$$

Then, \tilde{T}_i^2 can be written as

$$\begin{aligned} \tilde{T}_i^2 &= \left\{ \frac{n}{n-1}Q_i + \frac{n(n-2)}{(n-1)^3}(Q_iQ_j - Q_{i,j}^2) - \frac{2n}{(n-1)^2}Q_{i,j} + \frac{n}{(n-1)^3}Q_j \right\} \\ &\quad \times \left(1 + \frac{n-2}{(n-1)^2}Q_j \right)^{-1}, \end{aligned}$$

where

$$\begin{aligned} Q_i &= \left(1 + \frac{1}{n-2} \right) (\mathbf{X}_i - \bar{\mathbf{X}}_{(i,j)})' S_{(i,j)}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}_{(i,j)}), \\ Q_j &= \left(1 + \frac{1}{n-2} \right) (\mathbf{X}_j - \bar{\mathbf{X}}_{(i,j)})' S_{(i,j)}^{-1} (\mathbf{X}_j - \bar{\mathbf{X}}_{(i,j)}), \\ Q_{i,j} &= \left(1 + \frac{1}{n-2} \right) (\mathbf{X}_i - \bar{\mathbf{X}}_{(i,j)})' S_{(i,j)}^{-1} (\mathbf{X}_j - \bar{\mathbf{X}}_{(i,j)}), \end{aligned}$$

and

$$E\left(\hat{\beta}_{m,p}^2\right) = \frac{1}{n} E\left(T_i^{4m}\right) + \left(1 - \frac{1}{n}\right) E\left(T_i^{2m} T_j^{2m}\right).$$

Further, let

$$\begin{aligned} \bar{\mathbf{X}}_{(i,j)} &= \frac{1}{\sqrt{n-2}} \tilde{\mathbf{Y}}, \\ S_{(i,j)} &= I_p + \frac{1}{\sqrt{n-2}} \tilde{M}, \end{aligned}$$

then, \tilde{T}_i^{2m} can be expanded as

$$\tilde{T}_i^{2m} = (\mathbf{X}'_i \mathbf{X}_i)^m - \frac{1}{\sqrt{n}} m (\mathbf{X}'_i \mathbf{X}_i)^{m-1} f_1 + \frac{1}{n} m (\mathbf{X}'_i \mathbf{X}_i)^{m-2} e_1 + O_p\left(n^{-3/2}\right),$$

where

$$e_1 = \frac{m-1}{2} f_1^2 + \mathbf{X}'_i \mathbf{X}_i f_2 + (\mathbf{X}'_i \mathbf{X}_i)^2 - 2\mathbf{X}'_i \mathbf{X}_i \mathbf{X}'_i \mathbf{X}_j - \mathbf{X}'_i \mathbf{X}_i (\mathbf{X}'_i \mathbf{X}_j)^2,$$

and

$$\begin{aligned} f_1 &= 2\mathbf{X}'_i \tilde{\mathbf{Y}} + \mathbf{X}'_i \tilde{M} \mathbf{X}_i, \\ f_2 &= 2\mathbf{X}'_i \mathbf{X}_i + 2\mathbf{X}'_i \tilde{M} \tilde{\mathbf{Y}} + \mathbf{X}'_i \tilde{M}^2 \mathbf{X}_i + \tilde{\mathbf{Y}}' \tilde{\mathbf{Y}}. \end{aligned}$$

Substituting \tilde{T}_i^{2m} in T_i^{2m} , we can get the expansion of T_i^{4m} up to the order n^{-1} . Further calculating the expectation for the expansion of T_i^{4m} with respect to \mathbf{X}_i , \mathbf{Y} and M , and then doing for the expansion of $T_i^{2m} T_j^{2m}$ with respect to \mathbf{X}_i , \mathbf{X}_j , $\tilde{\mathbf{Y}}$ and \tilde{M} , we may come by the variance up to the order n^{-1} . Therefore, we have the following result.

THEOREM 2. *Suppose that $\hat{\mathcal{K}}_{(m)}$ is a consistent estimator of the $2m$ -th order moment parameter $\mathcal{K}_{(m)}$ with unknown Σ , which is given by (3.1). Then the asymptotic variance of $T = \sqrt{n}(\hat{\mathcal{K}}_{(m)} - \mathcal{K}_{(m)})$ is*

$$\begin{aligned} (3.6) \quad \sigma_T^2 &= \frac{\left(\frac{p}{2} + m\right)}{\binom{p}{2}_m} m (\mathcal{K}_{(2m)} + 1) + 3(m^2 - m - 2)(\kappa + 1) (\mathcal{K}_{(m)} + 1)^2 \\ &+ \frac{2^{-m+2} m^2}{p \binom{p}{2}_{m-2}} (\mathcal{K}_{(m)} + 1) (\kappa + 1) \{(p + 2)(\kappa + 1) - p\} \\ &+ (-m^2 + 3m + 1) (\mathcal{K}_{(m)} + 1)^2 - \frac{2m(p + 2m)}{p} (\mathcal{K}_{(m)} + 1) (\mathcal{K}_{(m+1)} + 1). \end{aligned}$$

This result is essentially the extension of that discussed by Seo and Toyama (1996). In the case when Σ is unknown, another estimator with the bias correction can be obtained as

$$(3.7) \quad \tilde{\mathcal{K}}_{(m)} = \hat{\mathcal{K}}_{(m)} - \frac{\hat{c}}{n},$$

where,

$$\begin{aligned} \hat{c} = & m \left(\hat{\mathcal{K}}_{(m-1)} + 1 \right) - m(p + 2m) \left(\hat{\mathcal{K}}_{(m+1)} + 1 \right) \\ & + (3 + mp + 2m)(\hat{\kappa} + 1) \left(\hat{\mathcal{K}}_{(m)} + 1 \right) - (m + 1) \left(\hat{\mathcal{K}}_{(m)} + 1 \right). \end{aligned}$$

Further, if the underlying distribution is normal then $\mathcal{K}_{(m)} = 0$, we have the variance of $\hat{\mathcal{K}}_{(m)}$ which is the same as that given by Mardia (1970).

4. Numerical calculation

This section investigates the bias and the MSE for the estimators of the general moment parameter by Monte Carlo simulation for some selected values of parameters. Computations are made for each case where the population covariance matrix Σ is unknown. For elliptical populations, we consider the following six types of distributions.

- E1 : Contaminated normal ($\omega = 0.1$, $\tau = 3$).
- E2 : Contaminated normal ($\omega = 0.4$, $\tau = 3$).
- E3 : Contaminated normal ($\omega = 0.7$, $\tau = 3$).
- E4 : Multivariate normal.
- E5 : Multivariate t ($\nu = 13$).
- E6 : Compound normal $U(1, 2)$,

where the random vector \mathbf{X} from E6 is the product of a normal vector \mathbf{Z} which has $N_p(\mathbf{0}, I_p)$ and the inverse of a random variable according to the uniform distribution on the interval $[1, 2]$. We note that the theoretical values of $\mathcal{K}_{(m)}$ are computed easily using the appropriate formula, i.e., for the contaminated normal distribution,

$$\mathcal{K}_{(m)} = \frac{1 + \omega(\tau^{2m} - 1)}{\{1 + \omega(\tau^2 - 1)\}^m} - 1,$$

for the multivariate t distribution with ν degrees of freedom is given by

$$\mathcal{K}_{(m)} = \frac{(\nu - 2)^m}{2^m \binom{\nu}{2 - m}_m} - 1, \quad \nu > 2m,$$

and for the compound normal distribution, $\mathcal{K}_{(m)} = 2^m(2^{-2m+1} - 1)/(-2m + 1) - 1$. The moment parameters of each distribution are obtained as follows Table 1, because we need them up to the sixth order for computing.

Table 1. Theoretical values of $\mathcal{K}_{(m)}$.

	κ	$\mathcal{K}_{(3)}$	$\mathcal{K}_{(4)}$	$\mathcal{K}_{(6)}$
E1	1.78	11.65	61.58	1561.52
E2	0.87	2.94	7.43	37.72
E3	0.31	0.77	1.42	3.50
E4	0	0	0	0
E5	0.22	0.92	3.22	169.42
E6	0.16	0.55	1.26	4.81

We first focus our concentration on the mean of T . Table 2 shows the approximate values with (3.5) for some parameters p and n when $m = 3$ and Σ is unknown. On the other hand, when we may assume that $\Sigma = I_p$ without any loss of generality, the sample mean for T based on 10,000 simulations is also obtained. It may be found from Table 2 that the simulation results nearly coincide with the approximate values for E4 and E6. But in the case of E5, the convergence is slow and the approximate expression (3.5) is not always precise. In the other cases, both values agree for a sufficiently large n .

Table 2. Simulation results for the mean of T when Σ is unknown.

$n \setminus p$	E1		E2		E3		E4		E5		E6		
	2	4	2	4	2	4	2	4	2	4	2	4	
100	AE	-92.50	-95.96	-9.37	-9.01	-3.17	-3.23	-1.30	-1.30	-7.59	-8.72	-3.46	-3.74
	SV	-58.32	-66.10	-6.95	-7.54	-1.98	-2.05	-1.20	-1.21	-4.66	-5.05	-2.39	-2.54
200	AE	-74.60	-86.24	-8.04	-8.49	-2.24	-2.28	-0.91	-0.91	-5.37	-6.16	-2.45	-2.64
	SV	-49.27	-57.98	-5.02	-5.55	-1.43	-1.50	-0.83	-0.83	-3.88	-4.41	-1.82	-1.93
500	AE	-47.18	-54.54	-5.08	-5.37	-1.41	-1.44	-0.58	-0.58	-3.39	-3.90	-1.54	-1.67
	SV	-36.46	-42.12	-3.48	-3.79	-0.97	-1.02	-0.50	-0.59	-3.44	-3.62	-1.06	-1.24
2000	AE	-23.59	-27.27	-2.54	-2.68	-0.71	-0.72	-0.29	-0.29	-1.69	-1.95	-0.77	-0.83
	SV	-21.13	-22.68	-1.89	-1.78	-0.48	-0.51	-0.27	-0.26	-2.02	-1.91	-0.50	-0.66

AE : Approximated values for the expectation of T .
 SV : Simulated values for the mean of T .

Next, Table 3 gives the bias for $\hat{\mathcal{K}}_{(3)}$ and another estimator $\tilde{\mathcal{K}}_{(3)}$ by (3.7) based on 10,000 simulations when Σ is unknown. The MSE is done by Table 4. If Σ is known, we make use of (3.4). It may be seen from simulation results in Tables 3 and 4 that the expectation of the estimators converges to the moment parameter when the sample size is large. Especially in normal case E4, it is noted that the modified estimator $\tilde{\mathcal{K}}_{(3)}$ rapidly approaches $\mathcal{K}_{(3)}$. The values obtained for the estimators were acceptable for a large n and improved as ω increased for the contaminated normal E1, E2 and E3, and when n increased for the compound normal E6. Also from Tables 3 and 4, we note that $\hat{\mathcal{K}}_{(3)}$ is underestimated for elliptical populations. The bias of $\tilde{\mathcal{K}}_{(3)}$ is actually smaller in magnitude than that of $\hat{\mathcal{K}}_{(3)}$. It may be noted that the size of p does not have much effect on the bias of $\hat{\mathcal{K}}_{(m)}$. These results are true for small sample sizes, not more than

$n = 50$, when Σ is known. As for the MSE, when Σ is unknown with large p , the MSE of $\tilde{\mathcal{K}}_{(m)}$ is about the same as that of $\hat{\mathcal{K}}_{(m)}$. It can be seen that the bias as well as the MSE for $\tilde{\mathcal{K}}_{(m)}$ is reduced when the sample size is large and the covariance matrix is unknown. As far as we can judge these results, $\tilde{\mathcal{K}}_{(m)}$ is better than $\hat{\mathcal{K}}_{(m)}$.

Table 3. Simulation results for the bias of the estimators when Σ is unknown.

p	n	E1		E2		E3		E4		E5		E6	
		$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$
2	100	-5.83	-2.73	-0.69	-0.13	-0.19	-0.04	-0.08	-0.01	-0.46	-0.15	-0.23	-0.01
	200	-3.48	-0.75	-0.35	-0.10	-0.10	-0.03	-0.04	-0.01	-0.27	-0.04	-0.12	-0.00
	500	-1.63	-0.07	-0.15	-0.05	-0.04	-0.01	-0.01	-0.00	-0.15	-0.04	-0.04	-0.01
	2000	-0.47	-0.02	-0.04	-0.01	-0.01	-0.00	-0.00	-0.00	-0.05	-0.00	-0.01	-0.00
4	100	-6.61	-3.35	-0.75	-0.11	-0.20	-0.03	-0.08	-0.01	-0.50	-0.18	-0.25	-0.01
	200	-4.10	-1.11	-0.39	-0.08	-0.10	-0.03	-0.04	-0.01	-0.31	-0.07	-0.13	-0.01
	500	-1.88	-0.08	-0.16	-0.04	-0.04	-0.01	-0.01	-0.00	-0.16	-0.04	-0.05	-0.01
	2000	-0.50	-0.05	-0.03	-0.01	-0.01	-0.00	-0.00	-0.00	-0.07	-0.03	-0.01	-0.00

Table 4. Simulation results for the MSE of the estimators when Σ is unknown.

p	n	E1		E2		E3		E4		E5		E6	
		$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$	$\hat{\mathcal{K}}_{(3)}$	$\tilde{\mathcal{K}}_{(3)}$
2	100	56.706	56.950	2.432	3.047	0.330	0.522	0.075	0.080	1.285	2.503	0.387	0.477
	200	42.093	42.164	1.604	1.785	0.201	0.205	0.042	0.050	1.144	1.438	0.239	0.278
	500	25.993	26.176	0.768	1.021	0.089	0.009	0.018	0.021	1.160	1.223	0.125	0.130
	2000	8.311	8.819	0.212	0.230	0.023	0.025	0.005	0.005	0.327	0.356	0.033	0.035
4	100	50.075	43.941	1.287	1.365	0.151	0.186	0.028	0.031	0.517	0.963	0.165	0.286
	200	26.174	26.222	0.704	1.025	0.080	0.096	0.014	0.021	0.433	0.872	0.099	0.103
	500	11.871	11.643	0.312	0.328	0.033	0.042	0.006	0.007	0.414	0.538	0.047	0.062
	2000	3.401	3.143	0.084	0.083	0.008	0.009	0.001	0.001	0.168	0.198	0.012	0.013

Furthermore, the asymptotic normality of T also enables us to easily construct the confidence intervals for $\mathcal{K}_{(m)}$. A confidence interval for $\mathcal{K}_{(m)}$ with confidence coefficient $1 - \alpha$ is approximately

$$(4.1) \quad \left[\hat{\mathcal{K}}_{(m)} - \sqrt{\frac{\hat{\sigma}_T^2}{n}} z_{\alpha/2}, \quad \hat{\mathcal{K}}_{(m)} + \sqrt{\frac{\hat{\sigma}_T^2}{n}} z_{\alpha/2} \right],$$

where $\hat{\sigma}_T^2$ is given by substituting $\hat{\mathcal{K}}_{(m)}$ for $\mathcal{K}_{(m)}$ in σ_T^2 gained as (3.3) in Theorem 1 or (3.6) in Theorem 2, because $\mathcal{K}_{(m)}$ is generally unknown. $z_{\alpha/2}$ is the two-tailed 100 α % point of the standard normal distribution. In the same way, we have another interval for $\mathcal{K}_{(m)}$ by means of $\tilde{\mathcal{K}}_{(m)}$, that is to say

$$(4.2) \quad \left[\tilde{\mathcal{K}}_{(m)} - \sqrt{\frac{\tilde{\sigma}_T^2}{n}} z_{\alpha/2}, \quad \tilde{\mathcal{K}}_{(m)} + \sqrt{\frac{\tilde{\sigma}_T^2}{n}} z_{\alpha/2} \right],$$

where $\tilde{\sigma}_T^2$ is done by substituting $\tilde{\mathcal{K}}_{(m)}$ for $\mathcal{K}_{(m)}$.

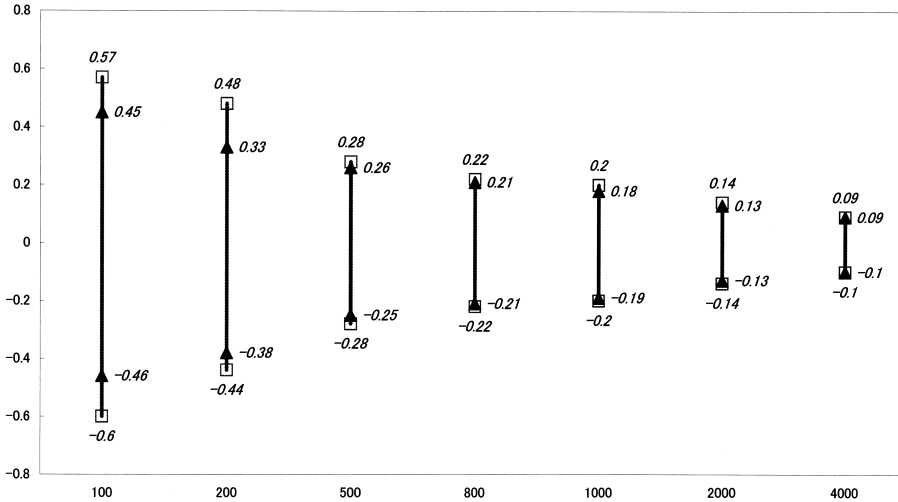


Figure 1. Confidence interval (E4). \square : $\hat{\mathcal{K}}_{(m)}$, \blacktriangle : $\tilde{\mathcal{K}}_{(m)}$

By way of illustration, we think over the normal case E4 when Σ is unknown. Figure 1 presents the confidence intervals for $\mathcal{K}_{(3)}$ with nominal confidence coefficient $\alpha = 0.95$. Those are the confidence limit with $\hat{\mathcal{K}}_{(3)}$ by (4.1) and with $\tilde{\mathcal{K}}_{(3)}$ by (4.2). It can be seen from those simulation results that the large sample size has the small range of the confidence interval. The confidence interval by means of $\tilde{\mathcal{K}}_{(3)}$ has a slightly small range by contrast with that of $\hat{\mathcal{K}}_{(3)}$. Moreover, the actual confidence coefficients of the interval in the circumstances are displayed in Table 5. It can turn out to be that both of the two values come closer to the nominal confidence coefficient $\alpha = 0.95$ when the sample size is large. The confidence coefficient with the help of $\hat{\mathcal{K}}_{(3)}$ is a little bigger than that of $\tilde{\mathcal{K}}_{(3)}$. Note that the largeness of p does not affect the intervals. In addition, when p is large, σ_T^2 decreases monotonously. In the other cases E1 to E6, the simulation results are shown in Table 6, and also then, we are able to see the results similar to E4.

Table 5. Actual confidence coefficients in E4.

n	100	200	500	800	1000	2000	4000
$\hat{\mathcal{K}}_{(3)}$	0.905	0.931	0.938	0.943	0.945	0.948	0.950
$\tilde{\mathcal{K}}_{(3)}$	0.885	0.928	0.938	0.942	0.944	0.948	0.950

The facts mentioned above may be applied to the case when Σ is known, but details are abbreviated in the present paper. Here, we cite only one or two instances. For the mean of T , it was seen that the simulation results almost agree with the approximate values (3.2) in each case. Also we found that the MSE of $\hat{\mathcal{K}}_{(m)}$ (or $\tilde{\mathcal{K}}_{(m)}$) is not as small as that in the case of an unknown Σ .

Table 6. Confidence interval for $\mathcal{K}_{(3)}$ when Σ is unknown.

$p = 2$	E1			E2			E3		
	lower	upper	α'	lower	upper	α'	lower	upper	α'
$n = 100 \hat{\mathcal{K}}_{(3)}$	-1.86	20.92	0.44	-0.15	4.85	0.80	-0.42	1.67	0.88
$\tilde{\mathcal{K}}_{(3)}$	2.01	20.41	0.40	1.40	4.80	0.62	-0.05	1.80	0.83
$n = 200 \hat{\mathcal{K}}_{(3)}$	-1.27	19.89	0.68	0.46	4.87	0.88	-0.12	1.53	0.91
$\tilde{\mathcal{K}}_{(3)}$	2.53	19.24	0.61	1.07	5.13	0.86	0.03	1.64	0.90
$n = 500 \hat{\mathcal{K}}_{(3)}$	1.29	19.71	0.86	1.26	4.36	0.92	0.18	1.30	0.93
$\tilde{\mathcal{K}}_{(3)}$	2.91	18.78	0.86	1.47	4.56	0.92	0.24	1.36	0.93
$n = 2000 \hat{\mathcal{K}}_{(3)}$	5.90	17.49	0.92	2.07	3.74	0.94	0.47	1.06	0.95
$\tilde{\mathcal{K}}_{(3)}$	6.28	17.08	0.92	2.16	3.80	0.95	0.49	1.07	0.95
$p = 2$	E4			E5			E6		
	lower	upper	α'	lower	upper	α'	lower	upper	α'
$n = 100 \hat{\mathcal{K}}_{(3)}$	-0.60	0.57	0.90	-1.33	2.33	0.14	-0.77	1.47	0.82
$\tilde{\mathcal{K}}_{(3)}$	-0.46	0.45	0.88	-1.24	2.88	0.28	-0.55	1.72	0.82
$n = 200 \hat{\mathcal{K}}_{(3)}$	-0.44	0.48	0.93	-1.62	2.96	0.43	-0.47	1.36	0.87
$\tilde{\mathcal{K}}_{(3)}$	-0.38	0.33	0.92	-1.86	3.66	0.51	-0.36	1.53	0.89
$n = 500 \hat{\mathcal{K}}_{(3)}$	-0.28	0.28	0.93	-1.21	2.76	0.56	-0.17	1.19	0.92
$\tilde{\mathcal{K}}_{(3)}$	-0.25	0.26	0.93	-1.40	3.19	0.63	-0.12	1.27	0.93
$n = 2000 \hat{\mathcal{K}}_{(3)}$	-0.14	0.14	0.95	-0.51	2.28	0.73	0.18	0.89	0.94
$\tilde{\mathcal{K}}_{(3)}$	-0.13	0.13	0.95	-0.57	2.24	0.76	0.19	0.91	0.94

α' : Actual confidence coefficients.

As a result, we conclude to suggest $\tilde{\mathcal{K}}_{(m)}$ as the better estimator of $\mathcal{K}_{(m)}$.

Acknowledgements

The authors wish to express their sincere gratitude to the referees for the comments that were very useful for the improvement of the paper.

REFERENCES

- Anderson, T. W. (1993). Nonnormal multivariate distributions: Inference based on elliptically contoured distributions, *Multivariate Analysis. Future Directions*, (ed. Rao, C. R.), 1–24, Elsevier Science Publishers.
- Berkane, M. and Bentler, P. M. (1986). Moments of elliptically distributed random variates, *Statist. Probab. Lett.*, **4**, 333–335.
- Berkane, M. and Bentler, P. M. (1987). Characterizing parameters of multivariate elliptical distributions, *Comm. Statist. Simulation Comput.*, **16**, 193–198.
- Cambanis, S., Huang, S. and Simons, G. (1981). On the theory of elliptically contoured distributions, *J. Multivariate Anal.*, **11**, 368–385.
- Hayakawa, T. and Puri, M. L. (1985). Asymptotic distributions of likelihood ratio criteria for testing latent roots and latent vectors of a covariance matrix under an elliptical population, *Biometrika*, **72**, 331–338.
- Henze, N. (1994). On Mardia's kurtosis test for multivariate normality, *Comm. Statist. Theory Methods*, **23**, 1031–1045.
- Kendall, M. G., Stuart, A. and Ord, J. K. (1987). *The Advanced Theory of Statistics, Vol.1*, 5th ed., Charles Griffin, London.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications, *Biometrika*, **57**, 519–530.

- Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, *Sankhyá*, **B36**, 115–128.
- Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, John Wiley, New York.
- Seo, T. and Toyama, T. (1996). On the estimation of kurtosis parameter in elliptical distributions, *J. Japan. Statist. Soc.*, **26**, 59–68.
- Wakaki, H. (1997). Asymptotic expansion of the joint distribution of sample mean vector and sample covariance matrix from an elliptical population, *Hiroshima Math. J.*, **27**, 295–305.