AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 8(3): 7346–7376.
DOI:10.3934/math.2023370
Received: 03 October 2022
Revised: 31 December 2022
Accepted: 03 January 2023
Published: 13 January 2023

Research article

Some results on multivariate measures of elliptical and skew-elliptical distributions: higher-order moments, skewness and kurtosis

Xueying Yu and Chuancun Yin*

School of Statistics and Data Science, Qufu Normal University, Qufu, Shandong 273165, China

* **Correspondence:** Email: ccyin@qfnu.edu.cn; Tel: 86+0537+4453221.

Abstract: The kurtosis and skewness of distributions are important measures that can describe the shape of a distribution, and there have been many results for symmetric distributions, but there are still many difficulties and challenges in the characterization of skew distributions. Based on the results of Mardia's and Song's kurtosis measures of elliptical distributions obtained by Zografos [1], we generalize the results and study some measures for elliptical and skew-elliptical distributions. We also derive the expressions of moments of skew-elliptical distributions in terms of the ones of skew-normals and take skew-*t*, skew-Pearson type VII and skew-Pearson type II distributions as examples.

Keywords: Mardia's measures; moments; multivariate elliptical distributions; multivariate skew-elliptical distributions; Song's measure of kurtosis

1. Introduction and motivation

Statisticians define the concepts of skewness and kurtosis in the case of univariate based on the standard normal distribution. Skewness is a measure of the degree of lopsidedness in the frequency distribution. Conversely, kurtosis is a measure of degree of tailedness in the frequency distribution. For symmetric distributions, all odd-order central moments (if any) are zero, so the first attempt to measure asymmetry is to compare the difference between the third-order central moments and zero. In principle, any other odd-order central moments can be used, or even their linear combination. Several optional asymmetry indices using modulo or quantiles are also available (see Arnold and Groeneveld [2], Avérous and Meste [3], Ekström and Jammalamadaka [4]). Similarly, for kurtosis, the kurtosis coefficient is obtained by using the fourth order central moment on the standard normal scale again. When dealing with multivariate distributions, the concepts of skewness as well as kurtosis measures are not uniquely defined – see for instance (based on moments like Mardia [5], based on union-intersetion principle of Roy [6] like Malkovich and Afifi [7], based on entropy like Song [8], etc.).

The family of multivariate elliptical distributions is a wide range of multivariate distribution family, which has many similar or even the same properties as multivariate normal. This family includes many well-known elliptical distributions, such as student-t, Pearson type II, Pearson type VII, symmetric Kotz type, among others. Studying the higher-order central moments of elliptical distributions is more conducive to our description of the shape of elliptical distributions. Mardia [5] proposed a common and popular method of measuring the skewness of multivariate distributions; Henze [9] tested the multivariate normality of the kurtosis measure proposed by Mardia; Azzalini and Dalla Valle [10] studied the multivariate skew-normal distribution and gave the moment generating function $M(t)$ for the case when $\boldsymbol{\mu} = \mathbf{0}$; Song [8] defined a general measure of the shape of distribution, that is, Song's measure, based on the entropy of λ introduced by Rényi [11]; Genton et al. [12] calculated the kurtosis and skewness of the skew-normal distribution when $\boldsymbol{\mu} \neq \mathbf{0}$.

In recent years, the results on the skewness and kurtosis of multivariate distributions are more abundant. Zografos [1] calculated the Mardia's measure and Song's measure of multivariate elliptical distributions, and found that the Mardia's measure is more sensitive to the center of the distribution while the Song's measure is more sensitive to the tail of the distribution. Kollo [13] compared skewness and kurtosis characteristics of multivariate distribution with results of Mardia [5] and used kurtosis matrix in ICA. Balakrishnan and Scarpa [14] calculated and compared different measures of skewness of multivariate skew-normal distributions. Tian et al. [15] introduced multivariate extended skew-normal distributions and discussed its quadratic forms. Kim and Kim [16] studied moments and quadratic forms of scale mixtures of skew-normal distributions. Jammalamadaka et al. [17] provided cumulant-based index of multivariate skewness and kurtosis and applied them to spherical, elliptical-symmetric and skew-symmetric families of multivariate distributions; Abdi et al. [18] constructed a new mixture family of multivariate normal distributions and derived the first four moments of it. Arellano-Valle and Azzalini [19] studied moments and Marida's skewness and kurtosis of continuous mixtures of multivariate normal distributions. Amiri and Balakrishnan [20] established Hessian and increasing-Hessian orderings of scale-shape mixtures of multivariate skew-normal distributions. Zuo and Yin [21] gave tail conditional expectation of generalized multivariate skew-elliptical distributions. This paper mainly studies some statistical measures of multivariate elliptical and multivariate skew-elliptical distributions, such as moments, Mardia's and Song's kurtosis measures.

The rest of this article is structured as follows. Section 2 generalizes the result of Mardia's kurtosis measure of elliptical distributions obtained by Zografos [1]. Section 3 introduces the skew-elliptical distributions and gives the expression of the first four moments of skew-elliptical distributions. In Section 4, the delta method is used to approximate Song's kurtosis measure of skew-elliptical distributions. In Section 5, Mardia's kurtosis measure of skew-elliptical distribution is obtained by transformation. Skew-t, skew-Pearson type VII and skew-Pearson type II distributions are taken as examples. In Section 6, we give a simple numerical analysis and consider the sample version for each of the measures considered as test statistics for the hypothesis of normal against the skew-normal distribution (SN). Finally, Section 7 gives a summary.

2. Mardia's kurtosis measure and its extension

This section extends Mardia's kurtosis measure to higher-order moments for a family of multivariate elliptical distributions. This family is a generalization of the family of multivariate normal distributions, including multivariate normal, multivariate t, multivariate Pearson type VII, multivariate logistic distributions and so on, which were discussed by Johnson [22]. The definitions and properties of some distributions in elliptical family were introduced in Fang et al. [23].

Definition 1. A p-dimensional random vector $\mathbf{X} = \left(X_1, ..., X_p\right)^T$ is said to have a multivariate elliptical distribution if its joint density function is given by

$$f(\mathbf{x}) = C_p |\mathbf{V}|^{-\frac{1}{2}} g\left[(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right], \tag{2.1}$$

where $\boldsymbol{\mu} \in \mathbb{R}^p$ is the location vector, $\mathbf{V}$ is the positive definite scale matrix, g is a non-negative, real valued function and satisfies $\int_0^\infty \omega^{p/2-1} g(\omega)\, \mathrm{d}\omega < \infty$, C_p is a normalized constant. In this case we shall write $\mathbf{X} \sim EC_p(\boldsymbol{\mu}, \mathbf{V}, g)$. Here g is called the density generator.

If $\mathbf{X} \sim EC_p(\boldsymbol{\mu}, \mathbf{V}, g)$, then the characteristic function of $\mathbf{X}$ has the form

$$\varphi(\mathbf{t}) = \exp\left(i\mathbf{t}^T\boldsymbol{\mu}\right)\psi\left(\mathbf{t}^T\mathbf{V}\mathbf{t}\right), \quad \mathbf{t} \in \mathbb{R}^p,$$

where ψ is a real valued function. The matrix $\mathbf{V}$ is proportional to the covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{X}$, i.e. $\boldsymbol{\Sigma} = Cov(\mathbf{X}) = -2\psi'(0)\,\mathbf{V}$.

There are many ways to define the measure of kurtosis for multivariate distributions. The most commonly used method is the measure introduced by Mardia [5] as follows:

$$\beta_{2,p} = E\left[(\mathbf{X}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu})\right]^2. \tag{2.2}$$

Zografos [1] calculated Mardia's kurtosis measure for multivariate elliptical distributions. Taking multivariate t, multivariate Pearson type VII, multivariate Pearson type II and symmetric Kotz type distributions as examples, the corresponding specific results were calculated as follows:
(1) Multivariate t-distribution:
The joint density function is given by

$$f(\mathbf{x}) = C_p |\mathbf{V}|^{-\frac{1}{2}} \left[1 + \frac{1}{\nu}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\nu+p)/2},$$

where $C_p = (\pi\nu)^{-\frac{p}{2}}\Gamma[(\nu+p)/2]/\Gamma(\nu/2)$, $-2\psi'(0) = \frac{\nu}{\nu-2}, \nu > 2$. The density generator is given by $g(\omega) = \left(1 + \frac{\omega}{\nu}\right)^{-(\nu+p)/2}$. The Mardia's kurtosis measure of the multivariate t-distribution is given by

$$\beta_{2,p} = \frac{p(p+2)(\nu-2)}{\nu-4},\ \nu > 4; \tag{2.3}$$

(2) Multivariate Pearson type VII distribution:
The joint density function is given by

$$f(\mathbf{x}) = C_p |\mathbf{V}|^{-\frac{1}{2}} \left[1 + (\mathbf{x}-\boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-m},$$

where $m > \frac{p}{2}$, $C_p = \pi^{-\frac{p}{2}}\Gamma(m)/\Gamma\left(m-\frac{p}{2}\right)$, $-2\psi'(0) = \frac{1}{2m-p-2}$. The density generator is given by $g(\omega) = (1+\omega)^{-m}$. The Mardia's kurtosis measure of the multivariate Pearson type VII distribution is given by

$$\beta_{2,p} = \frac{p\,(p+2)\,(2m-p-2)}{2m-p-4},\ m > \frac{p}{2}+2; \tag{2.4}$$

(3) Multivariate Pearson type II distribution:

The joint density function is given by

$$f(\mathbf{x}) = C_p|\mathbf{V}|^{-\frac{1}{2}}\left[1-(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^m,$$

where $\mathbf{x} \in S = \left\{\mathbf{x} \in \mathbb{R}^p : (\mathbf{x}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \le 1\right\}$, $m > -1$, $C_p = \pi^{-\frac{p}{2}}\ \Gamma[(p/2)+m+1]/\Gamma(m+1)$, $-2\psi'(0) = \frac{1}{2m+p+2}$ and $g(\omega) = (1-\omega)^m, 0 < \omega < 1$. The Mardia's kurtosis measure of the multivariate Pearson type II distribution is given by

$$\beta_{2,p} = \frac{p\,(p+2)\,(2m+p+2)}{2m+p+4},\ m > -1; \tag{2.5}$$

(4) Symmetric Kotz type distribution:

The joint density function is given by

$$\begin{aligned} f(\mathbf{x}) =& C_p|\mathbf{V}|^{-\frac{1}{2}}\left[(\boldsymbol{x}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^{m-1} \\ & \times \exp\left\{-r\left[(\boldsymbol{x}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^s\right\}, \end{aligned}$$

where $r > 0,\ s > 0,\ 2m+p > 2$,

$$-2\psi'(0) = \frac{\Gamma\left[(2m+p)/2s\right]r^{-1/s}}{p\Gamma\left[(2m+p-2)/2s\right]},$$

$$C_p = \frac{s\Gamma\left(p/2\right)r^{(2m+p-2)/2s}}{\pi^{p/2}\Gamma\left[(2m+p-2)/2s\right]}.$$

The density generator is given by $g(\omega) = \omega^{m-1}\exp\left(-r\omega^s\right)$. The Mardia's kurtosis measure of symmetric Kotz type multivariate distribution is given by

$$\beta_{2,p} = \frac{p^2\Gamma\left[(2m+p+2)/2s\right]\Gamma\left[(2m+p-2)/2s\right]}{\Gamma^2\left[(2m+p)/2s\right]},\ 2m+p > 2. \tag{2.6}$$

Inspired by Mardia and Zografos' work, we consider the following moments

$$\beta_{k,p} = E\left[(\mathbf{X}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})\right]^k,\ k \in \mathbb{N}^+ \tag{2.7}$$

for the family of elliptical distributions and take above four distributions as examples.

Theorem 2.1. *If* $\mathbf{X} \sim EC_p(\boldsymbol{\mu}, \mathbf{V}, g)$*, then*

$$\beta_{k,p} = \frac{\pi^{\frac{p}{2}}C_p}{\Gamma(p/2)\left[-2\psi'(0)\right]^k}\int_0^\infty \omega^{(2k+p)/2-1}g(\omega)\,\mathrm{d}\omega; \tag{2.8}$$

If, in addition, the density generator g *does not depend on the dimension* p*, then*

$$\beta_{k,p} = \frac{C_p\Gamma\left(\frac{p}{2}+k\right)}{\left[-2\psi'(0)\right]^k\pi^kC_{p+2k}\Gamma(p/2)}. \tag{2.9}$$

Proof. The proofs of results (2.8) and (2.9) are similar to that of Zografos [1]. □

Remark 1. In particular, when $k = 2$, we get the result in Zografos [1].

Using the above results, we can obtain the expressions of higher-order moments for specific members of the elliptical family. Next, we take multivariate t, multivariate Pearson type II and multivariate symmetric Kotz type distributions as examples.

Example 2.1 (Student-t distribution)

$$\beta_{k,p} = \frac{(\nu-2)^k \Gamma\left(\frac{p}{2}+k\right)\Gamma\left(\frac{\nu}{2}-k\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}, \ \nu > 2k.$$

In particular, in the case of $k = 2$, $\beta_{k,p}$ reduces to (2.3).

Example 2.2 (Pearson type II distribution)

$$\beta_{k,p} = \frac{(2m+p+2)^k \Gamma\left(\frac{p}{2}+m+1\right)\Gamma\left(\frac{p}{2}+k\right)}{\Gamma\left(\frac{p}{2}+m+k+1\right)\Gamma\left(\frac{p}{2}\right)}, \ m > -1.$$

In particular, in the case of $k = 2$, $\beta_{k,p}$ reduces to (2.5).

Example 2.3 (Symmetric Kotz type distribution)

$$\beta_{k,p} = \frac{p^k \Gamma^{k-1}\left(\frac{2m+p-2}{2s}\right)\Gamma\left(\frac{2m+p+2k-2}{2s}\right)}{\Gamma^k\left(\frac{2m+p}{2s}\right)}, \ 2m+p > 2.$$

In particular, in the case of $k = 2$, $\beta_{k,p}$ reduces to (2.6).

Next, letting $\boldsymbol{\mu} = (0,0)^T$ and

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix},$$

we calculated the values of $\beta_{k,p}$ at p=2 and k=2,3,4, for the above three examples. We also plotted the corresponding density function plots and density contour plots in Figures 1, 2, 3, and 4.

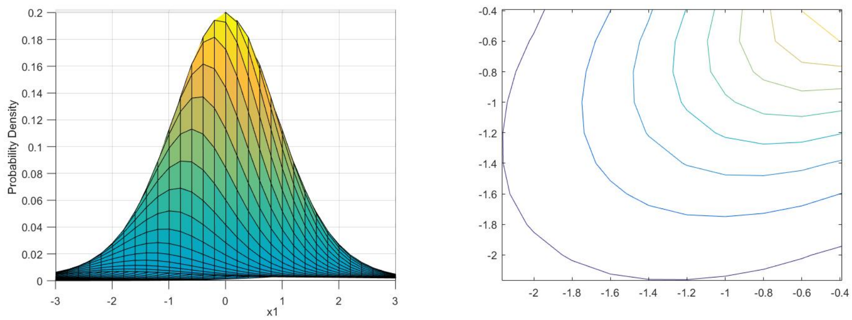

Figure 1. Student-t distribution with $\nu = 9$.

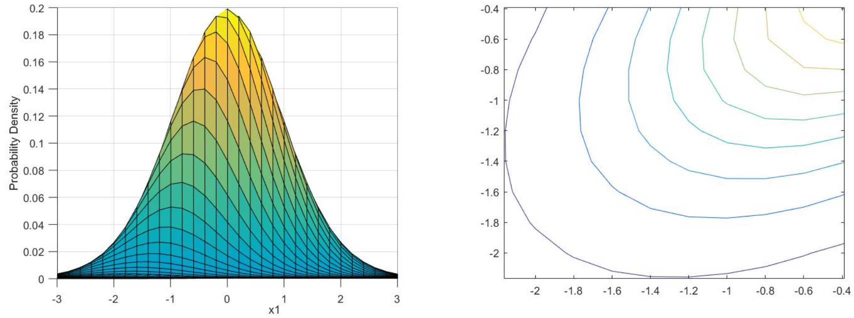

Figure 2. Student-t distribution with $\nu = 19$.

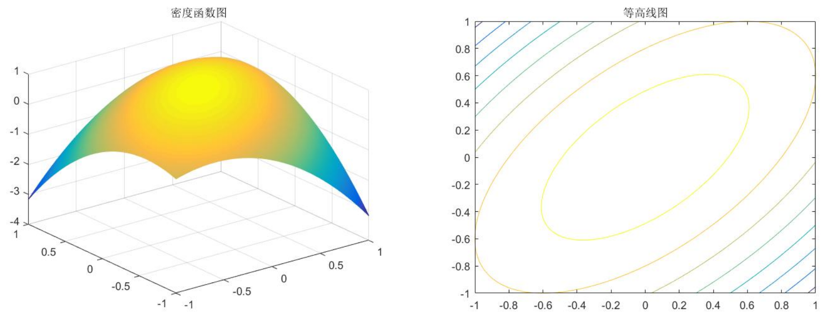

Figure 3. Pearson Type II distribution with $m = 1$.

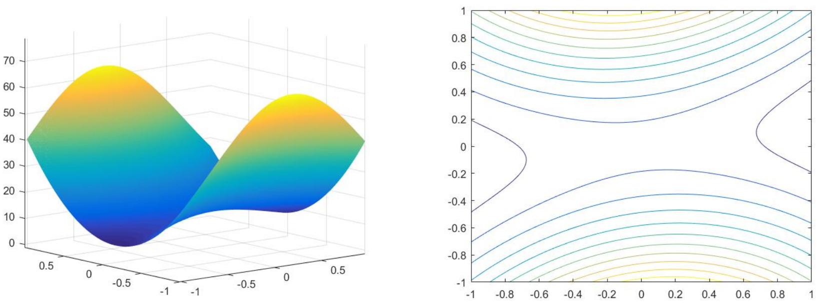

Figure 4. Symmetric Kotz type distribution with $m = 2$, $s = 1$, $r = 1/2$.

Table 1. Comparison of $\beta_{k,p}$ values of Student-t distribution with different degrees of freedom and comparison of $\beta_{k,p}$ values of different distributions.

Distribution	k	Parameter	$\beta_{k,p}$
Student-t	k=2	ν=9	11.2
	k=3	ν=9	156.8
	k=4	ν=9	8780.8
	k=2	ν=19	9.07
	k=3	ν=19	71.14
	k=4	ν=19	879.53
Pearson type II	k=2	m=1	6
	k=3	m=1	21.6
	k=4	m=1	86.4
Symmetric Kotz type	k=2	m=2, s=1	6
	k=3	m=2, s=1	24
	k=4	m=2, s=1	120

From Table 1, it can be seen that the values of $\beta_{k,p}$ all become larger as k increases, with the value of $\beta_{k,p}$ for the student-t distribution increasing steeply as k increases.

Figures 1 and 2 plot the density function plots and contour plots of the student-t distribution at the same mean vector and covariance matrices, with different degrees of freedom, and the corresponding values of $\beta_{k,p}$ are calculated in Table 1. It is very obvious from Table 1 that the value of $\beta_{k,p}$ decreases as ν increases, and the decrease is greater when k is larger, which is consistent with the properties of the student t-distribution. However, in the figures, we can see that the changes in the graphs and contours are not particularly pronounced, and the difference can be seen in the density contours between -1.2 and -1.4. Figures 3 and 4 plot the density function plots and contour plots for Pearson type II and symmetric Kotz type distributions for the same location, scale, skewness and kurtosis (Mardia indices), with the tail of symmetric Kotz type being relatively thicker. From the calculation results in Table 1, it can be seen that the difference of their $\beta_{k,p}$ values is larger as k increases.

Thus, it can be seen that the purpose of higher-order moments is to measure the degree of heavy tailing of a distribution and it applies to risk analysis. When the tails of a distribution are thicker, there is also a higher probability of events occurring for events that are further from the mean, and extreme events away from the centre of the distribution play a very important role. For example, a bankruptcy is more likely to come from a single extreme event. Therefore, the study of higher order moments is important.

3. Skew-elliptical distributions and their moments

Azzalini [24] introduced and studied the properties of unitary skew-normal distribution and its density function. Azzalini and Dalla Valle [10] put forward a general theory and probability properties of multivariate skew-normal distributions. Branco and Dey [25] generalized their results to multivariate skew-elliptical distributions. Some work moving in this direction can be found in recent review papers Azzalini [26] and Lee and McLachlan [27]. Azzalini and Capitanio [28] provided the main concepts and results of the skew-normal and related families, covering both the probability and the statistics sides of the subject in the univariate and multivariate settings. In the following, we adopt the notations in Branco and Dey [25].

Let $\mathbf{X}^* = \left(X_0, \mathbf{X}^T\right)^T \sim EC_{p+1}\left(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}, g^{(p+1)}\right)$, where $\mathbf{X} = \left(X_1, X_2, ..., X_p\right)^T$ is a p-dimensional random vector, $\boldsymbol{\mu}^* = \left(0, \boldsymbol{\mu}^T\right)^T, \boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_p)^T$, $g^{(p+1)} = C_{p+1} g(u; p+1)$, $g(u; p+1)$ and C_{p+1} are the generator function and the normalizing constant of the $p+1$ dimensional elliptic model, and the scale parameter matrix $\boldsymbol{\Sigma}$ has the following form

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \boldsymbol{\Omega} \end{pmatrix},$$

with $\boldsymbol{\delta} = \left(\delta_1, \delta_2, ..., \delta_p\right)^T$ and $\boldsymbol{\Omega}$ is the scale matrix of $\mathbf{X}$.

It is said that the random vector $\mathbf{Y} = \{\mathbf{X}|X_0 > 0\}$ has the skew-elliptical distribution and is denoted as $\mathbf{Y} \sim SE_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, if the density of random vector $\mathbf{X}^*$ exists and $P(\mathbf{X}^* = \mathbf{0}) = 0$, the density function of $\mathbf{Y}$ is as follows:

$$f_{\mathbf{Y}}(\mathbf{y}) = 2 f_{\hat{g}^{(p)}}(\mathbf{y}) F_{\tilde{g}_{q(\mathbf{y})}}\left(\boldsymbol{\alpha}^T(\mathbf{y} - \boldsymbol{\mu})\right), \tag{3.1}$$

where $f_{\hat{g}^{(p)}}(\cdot)$ is the density function of $EC_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \hat{g}^{(p)}\right)$, $\hat{g}^{(p)}$ is the marginal generator function, and $F_{\tilde{g}_{q(\mathbf{y})}}$ is the cumulative distribution function of $EC_1\left(0, 1, \tilde{g}_{q(\mathbf{y})}\right)$. Here,

$$\boldsymbol{\alpha}^T = \frac{\boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1}}{\left(1 - \boldsymbol{\delta}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\delta}\right)^{1/2}}, \tag{3.2}$$

$$\hat{g}^{(p)}(u) = \int_0^\infty r^{-\frac{1}{2}} g^{(p+1)}(r + u) \mathrm{d}r, \tag{3.3}$$

$$\tilde{g}_{q(\mathbf{y})}(u) = g^{(p+1)}\left[u + q(\mathbf{y})\right] / \hat{g}^{(p)}\left[q(\mathbf{y})\right], \tag{3.4}$$

$$q(\mathbf{y}) = (\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1} (\mathbf{y} - \boldsymbol{\mu}). \tag{3.5}$$

Family of skew-elliptical distributions includes many useful distributions, for examples, the multivariate skew-t, multivariate skew-Pearson type VII and multivariate skew-Pearson type II distributions.

Because it is complicated to directly find the moment generating function of skew-elliptical distribution, we will use the first four moments of skew-normal distribution obtained by Genton et al. [12] to get the first four moments of it through transformation and take multivariate skew-t, multivariate skew-Pearson type VII and multivariate skew-Pearson type II distributions as examples.

The following result is well known.

Lemma 3.1. *(Fang and Zhang [23]) Let* $\mathbf{X} \sim EC_p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g^{(p)}\right)$ *and* $\mathbf{B}_{p\times p}$ *is an invertible matrix, then*

$$\mathbf{B}\mathbf{X} \sim EC_p\left(\mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T; g^{(p)}\right).$$

Lemma 3.2. *(Gómez et al. [29]) p-dimensional random vector* $\mathbf{X} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ *if and only if*

$$\mathbf{X} \overset{d}{=} \boldsymbol{\mu} + \mathbf{A}^T R \mathbf{U}^{(p)},$$

where $\mathbf{A}$ *is a square matrix and* $\mathbf{A}^T\mathbf{A} = \boldsymbol{\Sigma}$, $\mathbf{U}^{(p)}$ *is a random vector with uniform distribution on the unit sphere in* $\mathbb{R}^p$, *R is an absolutely continuous non-negative random variable, and R is independent of* $\mathbf{U}^{(p)}$. *The density function of R is*

$$h(r) = \frac{2}{\int_0^\infty t^{\frac{p}{2}-1} g(t)\mathrm{d}t} r^{p-1} g\left(r^2\right), \; r \in (0, +\infty).$$

Inspired by Terdik [30] we get the following lemma.

Lemma 3.3. *Let* $\mathbf{U}^* = \left(U_1^*, \; \mathbf{U}_2^{*T}\right)^T \sim EC_{p+1}\left(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*, g^{(p+1)}\right)$, *where* $U_1^* = U_1$, $\mathbf{U}_2^* = \left(U_2, \cdots, U_{p+1}\right)^T$, $\boldsymbol{\mu}^* = \left(\mu_1^*, \; \boldsymbol{\mu}_2^{*T}\right)^T$ *and* $\mu_1^* = 0$, $\boldsymbol{\mu}_2^* = \left(\mu_1, \cdots, \mu_p\right)^T$,

$$\boldsymbol{\Sigma}^* = \begin{pmatrix} 1 & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & \boldsymbol{\Omega}^* \end{pmatrix},$$

where $\boldsymbol{\Omega}^*$ *is a p-dimensional symmetric matrix. Let* $Z_j = \delta_j |U_1| + \left(1 - \delta_j^2\right)^{\frac{1}{2}} \times \left(U_j - \mu_{j-1}\right) + \mu_{j-1}$; $-1 < \delta_j < 1$, $j = 2, 3, \cdots, p+1$, *then* $\mathbf{Z}_2^* = \left(Z_2, \cdots, Z_{p+1}\right)^T$ *has a density function equivalent to (3.1).*

Proof. Let $\mathbf{U}^{**} = \left(|U_1|, U_2, \cdots, U_{p+1}\right)^T \triangleq \left(U_1^{**}, \mathbf{U}_2^{**T}\right)^T$, then

$$\mathbf{U}^{**} \sim p_c^{-1} |\boldsymbol{\Sigma}^*|^{-\frac{1}{2}} g^{(p+1)} \left[(\mathbf{u}^{**} - \boldsymbol{\mu}^*)^T \boldsymbol{\Sigma}^{*\,-1} (\mathbf{u}^{**} - \boldsymbol{\mu}^*) \right],$$

where $p_c = P\left(U_1^{**} > 0\right) = \frac{1}{2}$, $g^{(p+1)} = C_{p+1} g(u; p+1)$.

Let $\mathbf{Z}^* = \left(Z_1, Z_2, \cdots, Z_{p+1}\right)^T \triangleq \left(Z_1^*, \mathbf{Z}_2^{*T}\right)^T$, and

$$\mathbf{B} \triangleq \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \delta_2 & \left(1-\delta_2^2\right)^{\frac{1}{2}} & 0 & \cdots & 0 \\ \delta_3 & 0 & \left(1-\delta_3^2\right)^{\frac{1}{2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \delta_{p+1} & 0 & 0 & \cdots & \left(1-\delta_{p+1}^2\right)^{\frac{1}{2}} \end{pmatrix} \triangleq \begin{pmatrix} 1 & \mathbf{0}_{1\times p} \\ \boldsymbol{\delta}_{p\times 1} & \boldsymbol{\Delta} \end{pmatrix},$$

where

$$\boldsymbol{\Delta} = diag\left\{ \left(1-\delta_2^2\right)^{\frac{1}{2}}, \cdots, \left(1-\delta_{p+1}^2\right)^{\frac{1}{2}} \right\},$$

then $\mathbf{Z}^* = \mathbf{B}(\mathbf{U}^{**} - \boldsymbol{\mu}^*) + \boldsymbol{\mu}^*$.

From Lemma 3.1, we can get

$$\begin{aligned} \mathbf{Z}^* \sim & p_c^{-1} |\boldsymbol{\Sigma}^*|^{-\frac{1}{2}} |\mathbf{B}|^{-1} g^{(p+1)} \left[(\mathbf{z}^* - \boldsymbol{\mu}^*)^T \left(\mathbf{B}^{-1}\right)^T \boldsymbol{\Sigma}^{*\,-1} \mathbf{B}^{-1} (\mathbf{z}^* - \boldsymbol{\mu}^*) \right] \\ = & p_c^{-1} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} g^{(p+1)} \left[(\mathbf{z}^* - \boldsymbol{\mu}^*)^T \boldsymbol{\Sigma}^{-1} (\mathbf{z}^* - \boldsymbol{\mu}^*) \right], \end{aligned}$$

where

$$\mathbf{\Sigma} \triangleq \mathbf{B}\mathbf{\Sigma}^*\mathbf{B}^T = \begin{pmatrix} 1 & \mathbf{0}_{1\times p} \\ \boldsymbol{\delta}_{p\times 1} & \mathbf{\Delta} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & \mathbf{\Omega}^* \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{\delta}_{1\times p}^T \\ \mathbf{0}_{p\times 1} & \mathbf{\Delta} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \boldsymbol{\delta}\boldsymbol{\delta}^T + \mathbf{\Delta}\mathbf{\Omega}^*\mathbf{\Delta} \end{pmatrix}.$$

Let $t = \mathbf{\Sigma}_{11.2}^{-\frac{1}{2}}\left(z_1^* - \boldsymbol{\mu}_{1.2}^*\right)$, the marginal density of $\mathbf{Z}_2^*$ is

$$\begin{aligned}
\mathbf{Z}_2^* \sim & p_c^{-1}|\mathbf{\Sigma}|^{-\frac{1}{2}} \int_{z_1^*>0} g^{(p+1)} \left[(\mathbf{z}^* - \boldsymbol{\mu}^*)^T \mathbf{\Sigma}^{-1} (\mathbf{z}^* - \boldsymbol{\mu}^*) \right] \mathrm{d}z_1^* \\
=& p_c^{-1} f_{\hat{g}^{(p)}} (\mathbf{z}_2^*) \int_{z_1^*>0} f_{\tilde{g}_{q(\mathbf{z}_2^*)}} (z_1^* | \mathbf{z}_2^*) \, \mathrm{d}z_1^* \\
=& 2 f_{\hat{g}^{(p)}} (\mathbf{z}_2^*) \int_{z_1^*>0} |\mathbf{\Sigma}_{11.2}|^{-\frac{1}{2}} \tilde{g}_{q(\mathbf{z}_2^*)} \left[(z_1^* - \boldsymbol{\mu}_{1.2}^*)^2 \mathbf{\Sigma}_{11.2}^{-1} (z_1^* - \boldsymbol{\mu}_{1.2}^*) \right] \mathrm{d}z_1^* \\
=& 2 f_{\hat{g}^{(p)}} (\mathbf{z}_2^*) \int_{t<\mathbf{\Sigma}_{11.2}^{-\frac{1}{2}}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{z}_2^*-\boldsymbol{\mu}_2^*)} \tilde{g}_{q(\mathbf{z}_2^*)} \left(t^2\right) \mathrm{d}t \\
=& 2 f_{\hat{g}^{(p)}} (\mathbf{z}_2^*) F_{\tilde{g}_{q(\mathbf{z}_2^*)}} \left[\boldsymbol{\alpha}^T (\mathbf{z}_2^* - \boldsymbol{\mu}_2^*) \right],
\end{aligned}$$

where $\boldsymbol{\alpha}^T = \mathbf{\Sigma}_{11.2}^{-\frac{1}{2}}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}$. This formula is equivalent to (3.1).
This completes the proof of Lemma 3.3. □

From Lemma 3.3, we can obtain a lemma as follows:

Lemma 3.4. *If* $\mathbf{Y} \sim SE_p\left(\boldsymbol{\mu}, \mathbf{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, *then*

$$\mathbf{Y} \overset{d}{=} R\left(\boldsymbol{\delta}|U^{(1)}| + \mathbf{\Delta}\mathbf{A}^{*T}\mathbf{U}^{(p)}\right) + \boldsymbol{\mu},$$

where R is an absolutely continuous non-negative random variable, $\left(U^{(1)}, \mathbf{U}^{(p)T}\right)^T$ has the uniform distribution on the unit sphere in $\mathbb{R}^{p+1}$ and independent of R, $\mathbf{A}^{*T}\mathbf{A}^* = \mathbf{\Omega}$, $\mathbf{\Delta} = diag\left\{\left(1-\delta_2^2\right)^{\frac{1}{2}}, \cdots, \left(1-\delta_{p+1}^2\right)^{\frac{1}{2}}\right\}$.

Proof. By Lemma 3.2, we obtain

$$\left(U_1^*, \mathbf{U}_2^{*T}\right)^T \overset{d}{=} \boldsymbol{\mu}^* + R\begin{pmatrix} 1 & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & \mathbf{A}^{*T} \end{pmatrix} \mathbf{U}^{(p+1)},$$

where $\mathbf{U}^{(p+1)} = \left(U^{(1)}, \mathbf{U}^{(p)T}\right)^T$ has the uniform distribution on the unit sphere in $\mathbb{R}^{p+1}$, $\boldsymbol{\mu}^* = \left(0, \boldsymbol{\mu}^T\right)$, $\mathbf{A}^{*T}\mathbf{A}^* = \mathbf{\Omega}$ and R is independent of $\mathbf{U}^{(p+1)}$. The density of R is

$$h_R(r) = \frac{2\pi^{\frac{p+1}{2}}}{\Gamma\left(\frac{p+1}{2}\right)} r^p g^{(p+1)}\left(r^2\right),$$

where $g^{(p+1)}\left(r^2\right) = C_{p+1}\left(r^2; p+1\right)$. From Lemma 3.3, we have

$$\begin{aligned}\mathbf{Y} &\overset{d}{=} \boldsymbol{\delta}|U_1^*| + \boldsymbol{\Delta}\left(\mathbf{U}_2^* - \boldsymbol{\mu}\right) + \boldsymbol{\mu} \\ &\overset{d}{=} \boldsymbol{\delta}|RU^{(1)}| + \boldsymbol{\Delta}\left(\boldsymbol{\mu} + R\mathbf{A}^{*T}\mathbf{U}^{(p)} - \boldsymbol{\mu}\right) + \boldsymbol{\mu} \\ &= R\left(\boldsymbol{\delta}|U^{(1)}| + \boldsymbol{\Delta}\mathbf{A}^{*T}\mathbf{U}^{(p)}\right) + \boldsymbol{\mu}.\end{aligned} \tag{3.6}$$

□

Remark 2. If $\mathbf{X} \sim SN_p\left(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\delta}\right)$, we can obtain

$$\mathbf{X} \overset{d}{=} R_0\left(\boldsymbol{\delta}|U^{(1)}| + \boldsymbol{\Delta}\mathbf{A}^{*T}\mathbf{U}^{(p)}\right), \tag{3.7}$$

where $R_0 \sim \chi_{p+1}$. See, for example, Genton [31].

Lemma 3.5. *(Genton et al. [12]) If $\mathbf{Z} \sim SN_p(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\delta})$, then the first four moments of $\mathbf{Z}$ are*

(1) $M_1 = \sqrt{\frac{2}{\pi}}\boldsymbol{\delta}$; (2) $M_2 = \boldsymbol{\Omega}$;

(3) $M_3 = \sqrt{\frac{2}{\pi}}\left[\boldsymbol{\delta} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T + \left(\mathbf{I}_n \otimes \boldsymbol{\delta}\right)\boldsymbol{\Omega} - \left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right)\left(\boldsymbol{\delta} \otimes \boldsymbol{\delta}^T\right)\right]$;

(4) $M_4 = \left[\left(\mathbf{I}_{p^2} + \mathbf{U}_{p,p}\right)\boldsymbol{\Omega} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega})\left(vec(\boldsymbol{\Omega})\right)^T\right]$.

Theorem 3.1. *If $\mathbf{Y} \sim SE_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, then the first four moments of $\mathbf{Y}$ are expressed as follows:*

(1) $M_1 = \boldsymbol{\mu} + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}\boldsymbol{\delta}$;

(2) $M_2 = \boldsymbol{\mu}\boldsymbol{\mu}^T + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}\left(\boldsymbol{\mu}\boldsymbol{\delta}^T + \boldsymbol{\delta}\boldsymbol{\mu}^T\right) + \frac{E\left(R^2\right)}{E\left(R_0^2\right)}\boldsymbol{\Omega}$;

(3) $$\begin{aligned}M_3 = {} & \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}[\boldsymbol{\delta} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} + \boldsymbol{\mu} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\mu} \\ & + \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\delta}] + \frac{E\left(R^2\right)}{E\left(R_0^2\right)}\left[\boldsymbol{\Omega} \otimes \boldsymbol{\mu} + \boldsymbol{\mu} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega}) \otimes \boldsymbol{\mu}^T\right] \\ & + \frac{E(R^3)}{E(R_0^3)}\sqrt{\frac{2}{\pi}}\left[\boldsymbol{\delta} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T + \left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right)\boldsymbol{\Omega} - \left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right)\left(\boldsymbol{\delta} \otimes \boldsymbol{\delta}^T\right)\right];\end{aligned}$$

(4) $$\begin{aligned}M_4 = {} & \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}G_1 + \frac{E(R^2)}{E(R_0^2)}G_2 + \frac{E(R^3)}{E(R_0^3)}\sqrt{\frac{2}{\pi}}G_3 \\ & + \frac{E(R^4)}{E(R_0^4)}G_4,\end{aligned}$$

where

$$
\begin{aligned}
G_1 &= \boldsymbol{\delta} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T + \boldsymbol{\mu} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T + \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\delta} \otimes \boldsymbol{\mu}^T \\
&\quad + \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} \otimes \boldsymbol{\delta}^T; \\
G_2 &= \boldsymbol{\Omega} \otimes \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T + \boldsymbol{\mu} \otimes \boldsymbol{\Omega} \otimes \boldsymbol{\mu}^T + vec(\boldsymbol{\Omega}) \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu}^T + \boldsymbol{\mu}^T \otimes \boldsymbol{\Omega} \otimes \boldsymbol{\mu} \\
&\quad + \boldsymbol{\mu} \otimes \boldsymbol{\mu} \otimes (vec(\boldsymbol{\Omega}))^T + \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\Omega}; \\
G_3 &= \boldsymbol{\delta} \otimes \boldsymbol{\Omega} \otimes \boldsymbol{\mu}^T + vec(\boldsymbol{\Omega}) \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\mu}^T + \left(\left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right) \boldsymbol{\Omega}\right) \otimes \boldsymbol{\mu}^T + \boldsymbol{\delta}^T \otimes \boldsymbol{\Omega} \otimes \boldsymbol{\mu} \\
&\quad + \boldsymbol{\delta} \otimes (vec(\boldsymbol{\Omega}))^T \otimes \boldsymbol{\mu} + \left(\boldsymbol{\Omega}\left(\mathbf{I}_p \otimes \boldsymbol{\delta}^T\right)\right) \otimes \boldsymbol{\mu} + \boldsymbol{\mu}^T \otimes \boldsymbol{\delta} \otimes \boldsymbol{\Omega} \\
&\quad + \boldsymbol{\mu}^T \otimes \left(vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T\right) + \boldsymbol{\mu}^T \otimes \left(\left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right) \boldsymbol{\Omega}\right) + \boldsymbol{\mu} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\Omega} + \boldsymbol{\mu} \otimes \boldsymbol{\delta} \otimes (vec(\boldsymbol{\Omega}))^T \\
&\quad + \boldsymbol{\mu} \otimes \left(\boldsymbol{\Omega}\left(\mathbf{I}_p \otimes \boldsymbol{\delta}^T\right)\right) - \boldsymbol{\delta} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\delta} \otimes \boldsymbol{\mu}^T - \boldsymbol{\delta}^T \otimes \boldsymbol{\delta} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\mu} \\
&\quad - \boldsymbol{\mu}^T \otimes \boldsymbol{\delta} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\delta} - \boldsymbol{\mu} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\delta} \otimes \boldsymbol{\delta}^T; \\
G_4 &= \left(\mathbf{I}_{p^2} + \mathbf{U}_{p,p}\right) \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega})\, (vec(\boldsymbol{\Omega}))^T .
\end{aligned}
$$

Proof. From Lemma 3.4 and Remark 2, we have known that

$$\mathbf{Y} \overset{d}{=} R\mathbf{M} + \boldsymbol{\mu}, \ \ \mathbf{X} \overset{d}{=} R_0\mathbf{M}, \ \mathbf{X} \sim SN_p(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\delta}),$$

where $\mathbf{M} = \boldsymbol{\delta}|U^{(1)}| + \boldsymbol{\Delta}\mathbf{A}^{*T}\mathbf{U}^{(p)}$, then

$$E(\mathbf{Y}) = E(R)E(\mathbf{M}) + \boldsymbol{\mu}; \ \ E(\mathbf{X}) = E(R_0)E(\mathbf{M}).$$

So we obtain

$$
\begin{aligned}
E(\mathbf{M}) &= \frac{E(\mathbf{X})}{E(R_0)}; \\
E(\mathbf{Y}) &= \boldsymbol{\mu} + \frac{E(R)}{E(R_0)} E(\mathbf{X}).
\end{aligned}
$$

In the same way, we have

$$
\begin{aligned}
E\left(\mathbf{M}\mathbf{M}^T\right) &= \frac{E\left(\mathbf{X}\mathbf{X}^T\right)}{E\left(R_0^2\right)}; \ \ E\left(\mathbf{M} \otimes \mathbf{M}^T \otimes \mathbf{M}\right) = \frac{E\left(\mathbf{X} \otimes \mathbf{X}^T \otimes \mathbf{X}\right)}{E\left(R_0^3\right)}; \\
E\left(\mathbf{M} \otimes \mathbf{M}^T \otimes \mathbf{M} \otimes \mathbf{M}^T\right) &= \frac{E\left(\mathbf{X} \otimes \mathbf{X}^T \otimes \mathbf{X} \otimes \mathbf{X}^T\right)}{E\left(R_0^4\right)}.
\end{aligned}
$$

According to Lemma 3.5 and using above relationship, we obtain

$$
\begin{aligned}
M_1 &= E(\mathbf{Y}) = \boldsymbol{\mu} + \frac{E(R)}{E(R_0)} E(\mathbf{X}) = \boldsymbol{\mu} + \frac{E(R)}{E(R_0)} \sqrt{\frac{2}{\pi}} \boldsymbol{\delta}; \\
M_2 &= E\left[(\boldsymbol{\mu} + R\mathbf{M})(\boldsymbol{\mu} + R\mathbf{M})^T\right] \\
&= E\left(\boldsymbol{\mu}\boldsymbol{\mu}^T + R^2\mathbf{M}\mathbf{M}^T + R\boldsymbol{\mu}\mathbf{M}^T + R\mathbf{M}\boldsymbol{\mu}^T\right) \\
&= \boldsymbol{\mu}\boldsymbol{\mu}^T + \frac{E(R)}{E(R_0)} \sqrt{\frac{2}{\pi}} \left(\boldsymbol{\mu}\boldsymbol{\delta}^T + \boldsymbol{\delta}\boldsymbol{\mu}^T\right) + \frac{E\left(R^2\right)}{E\left(R_0^2\right)} \boldsymbol{\Omega};
\end{aligned}
$$

$$\begin{aligned}
M_2 &= E\left[(\boldsymbol{\mu}+R\mathbf{M})(\boldsymbol{\mu}+R\mathbf{M})^T\right]\\
&= E\left(\boldsymbol{\mu}\boldsymbol{\mu}^T + R^2\mathbf{M}\mathbf{M}^T + R\boldsymbol{\mu}\mathbf{M}^T + R\mathbf{M}\boldsymbol{\mu}^T\right)\\
&= \boldsymbol{\mu}\boldsymbol{\mu}^T + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}\left(\boldsymbol{\mu}\boldsymbol{\delta}^T + \boldsymbol{\delta}\boldsymbol{\mu}^T\right) + \frac{E\left(R^2\right)}{E\left(R_0^2\right)}\boldsymbol{\Omega};\\
M_3 &= E\left(\mathbf{Y}\otimes\mathbf{Y}^T\otimes\mathbf{Y}\right) = E\left[(\boldsymbol{\mu}+R\mathbf{M})\otimes(\boldsymbol{\mu}+R\mathbf{M})^T\otimes(\boldsymbol{\mu}+R\mathbf{M})\right]\\
&= E[\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu} + R(\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\mathbf{M}) + R^2(\mathbf{M}\otimes\mathbf{M}^T\otimes\boldsymbol{\mu})\\
&\quad + R^3(\mathbf{M}\otimes\mathbf{M}^T\otimes\mathbf{M}) + R(\boldsymbol{\mu}\otimes\mathbf{M}^T\otimes\boldsymbol{\mu}) + R^2(\boldsymbol{\mu}\otimes\mathbf{M}^T\otimes\mathbf{M})\\
&\quad + R(\mathbf{M}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}) + R^2(\mathbf{M}\otimes\boldsymbol{\mu}^T\otimes\mathbf{M})]\\
&= \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu} + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}\left[\boldsymbol{\delta}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu} + \boldsymbol{\mu}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\mu} + \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\delta}\right]\\
&\quad + \frac{E(R^2)}{E(R_0^2)}\left[\boldsymbol{\Omega}\otimes\boldsymbol{\mu} + \boldsymbol{\mu}\otimes\boldsymbol{\Omega} + vec(\boldsymbol{\Omega})\otimes\boldsymbol{\mu}^T\right] + \frac{E(R^3)}{E(R_0^3)}\sqrt{\frac{2}{\pi}}[\boldsymbol{\delta}\otimes\boldsymbol{\Omega}\\
&\quad + vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T + (\mathbf{I}_p\otimes\boldsymbol{\delta})\boldsymbol{\Omega} - \left(\mathbf{I}_p\otimes\boldsymbol{\delta}\right)\left(\boldsymbol{\delta}\otimes\boldsymbol{\delta}^T\right)];\\
M_4 &= E[(\boldsymbol{\mu}+R\mathbf{M})\otimes(\boldsymbol{\mu}+R\mathbf{M})^T\otimes(\boldsymbol{\mu}+R\mathbf{M})\otimes(\boldsymbol{\mu}+R\mathbf{M})^T]\\
&= \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T + \frac{E(R)}{E(R_0)}\sqrt{\frac{2}{\pi}}[\boldsymbol{\delta}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T + \boldsymbol{\mu}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\\
&\quad + \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\delta}\otimes\boldsymbol{\mu}^T + \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\delta}^T] + \frac{E\left(R^2\right)}{E\left(R_0^2\right)}[\boldsymbol{\Omega}\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\\
&\quad + \boldsymbol{\mu}\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\mu}^T + vec(\boldsymbol{\Omega})\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}^T + \boldsymbol{\mu}^T\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\mu}\\
&\quad + \boldsymbol{\mu}\otimes\boldsymbol{\mu}\otimes(vec(\boldsymbol{\Omega}))^T + \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\Omega}] + \frac{E\left(R^3\right)}{E\left(R_0^3\right)}\sqrt{\frac{2}{\pi}}[\boldsymbol{\delta}\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\mu}^T\\
&\quad + vec(\boldsymbol{\Omega})\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\mu}^T + \left((\mathbf{I}_p\otimes\boldsymbol{\delta}\right)\boldsymbol{\Omega})\otimes\boldsymbol{\mu}^T + \boldsymbol{\delta}^T\otimes\boldsymbol{\Omega}\otimes\boldsymbol{\mu}\\
&\quad + \boldsymbol{\delta}\otimes(vec(\boldsymbol{\Omega}))^T\otimes\boldsymbol{\mu} + (\boldsymbol{\Omega}(\mathbf{I}_p\otimes\boldsymbol{\delta}^T))\otimes\boldsymbol{\mu} + \boldsymbol{\mu}^T\otimes\boldsymbol{\delta}\otimes\boldsymbol{\Omega}\\
&\quad + \boldsymbol{\mu}^T\otimes\left(vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T\right) + \boldsymbol{\mu}^T\otimes\left((\mathbf{I}_p\otimes\boldsymbol{\delta})\boldsymbol{\Omega}\right) + \boldsymbol{\mu}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\Omega}\\
&\quad + \boldsymbol{\mu}\otimes\boldsymbol{\delta}\otimes(vec(\boldsymbol{\Omega}))^T + \boldsymbol{\mu}\otimes\left(\boldsymbol{\Omega}(\mathbf{I}_p\otimes\boldsymbol{\delta}^T)\right) - \boldsymbol{\delta}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\delta}\otimes\boldsymbol{\mu}^T\\
&\quad - \boldsymbol{\delta}^T\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\mu} - \boldsymbol{\mu}^T\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\delta} - \boldsymbol{\mu}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\delta}\otimes\boldsymbol{\delta}^T]\\
&\quad + \frac{E\left(R^4\right)}{E\left(R_0^4\right)}\left[(\mathbf{I}_{p^2} + \mathbf{U}_{p,p})\boldsymbol{\Omega}\otimes\boldsymbol{\Omega} + vec(\boldsymbol{\Omega})(vec(\boldsymbol{\Omega}))^T\right].
\end{aligned}$$

□

Remark 3. If $R_0^2 \sim \chi_{p+1}^2$ and χ_{p+1}^2 is a central chi square distribution with $p+1$ degrees of freedom, then

$$E(R_0) = \frac{\sqrt{2}\Gamma\left(\frac{p}{2}+1\right)}{\Gamma\left(\frac{p+1}{2}\right)}; \quad E\left(R_0^2\right) = p+1;$$

$$E\left(R_0^3\right) = \frac{2^{\frac{3}{2}}\Gamma\left(\frac{p}{2}+2\right)}{\Gamma\left(\frac{p+1}{2}\right)}; \quad E\left(R_0^4\right) = (p+1)(p+3).$$

Using above results, we give the first four moments of multivariate skew-t (St), multivariate skew-Pearson type VII (SPVII) and multivariate skew-Pearson type II (SPII) distributions, respectively.

Example 3.1 If $\mathbf{Y} = \{\mathbf{X}|X_0 > 0\} \sim St\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$ and its joint density function is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = 2|\Omega|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma(\nu/2)(\pi\nu)^{p/2}} [1 + \frac{(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1} (\mathbf{y}-\boldsymbol{\mu})}{\nu}]^{-\frac{\nu+p}{2}}$$
$$\times \int_{-\infty}^{\alpha^T(\mathbf{y}-\boldsymbol{\mu})} \frac{\Gamma[(\nu+p+1)/2][1+\frac{(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\nu}]^{(\nu+p)/2}}{\sqrt{\pi\nu}\Gamma[(\nu+p)/2][1+\frac{(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})+r^2}{\nu}]^{(\nu+p+1)/2}} \mathrm{d}r,$$

with

$$g^{(p+1)}(u) = \frac{\Gamma[(\nu+p+1)/2]}{\Gamma(\nu/2)(\pi\nu)^{(p+1)/2}} \left(1+\frac{u}{\nu}\right)^{-\frac{\nu+p+1}{2}},$$

then the first four moments of $\mathbf{Y}$ are expressed as follows:

(1) $M_1 = \boldsymbol{\mu} + \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \boldsymbol{\delta}, \ \nu > 1;$

(2) $M_2 = \boldsymbol{\mu}\boldsymbol{\mu}^T + \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\boldsymbol{\mu}\boldsymbol{\delta}^T + \boldsymbol{\delta}\boldsymbol{\mu}^T\right) + \frac{\nu}{\nu-2}\boldsymbol{\Omega}, \ \nu > 2;$

(3)
$$M_3 = \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} + \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} [\boldsymbol{\delta} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} + \boldsymbol{\mu} \otimes \boldsymbol{\delta}^T \otimes \boldsymbol{\mu}$$
$$+ \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\delta}] + \frac{\nu}{\nu-2}[\boldsymbol{\Omega} \otimes \boldsymbol{\mu} + \boldsymbol{\mu} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega}) \otimes \boldsymbol{\mu}^T] + \left(\frac{\nu}{2}\right)^{\frac{3}{2}}$$
$$\times \sqrt{\frac{2}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}-1\right)}{\Gamma\left(\frac{\nu}{2}\right)} [\boldsymbol{\delta} \otimes \boldsymbol{\Omega} + vec(\boldsymbol{\Omega}) \boldsymbol{\delta}^T + \left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right) \boldsymbol{\Omega}$$
$$- \left(\mathbf{I}_p \otimes \boldsymbol{\delta}\right)\left(\boldsymbol{\delta} \otimes \boldsymbol{\delta}^T\right)], \ \nu > 3;$$

(4)
$$M_4 = \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T \otimes \boldsymbol{\mu} \otimes \boldsymbol{\mu}^T + \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} G_1 + \frac{\nu}{\nu-2} G_2 + \left(\frac{\nu}{2}\right)^{\frac{3}{2}} \frac{\Gamma\left(\frac{\nu-1}{2}-1\right)}{\Gamma\left(\frac{\nu}{2}\right)}$$
$$\times \sqrt{\frac{2}{\pi}} G_3 + \frac{\nu^2}{(\nu-2)(\nu-4)} G_4, \ \nu > 4.$$

Proof. See the Appendix. □

Example 3.2 If $\mathbf{Y} = \{\mathbf{X}|X_0 > 0\} \sim SPVII\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$ and its joint density function is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = 2|\boldsymbol{\Omega}|^{-\frac{1}{2}} \frac{\Gamma\left(m-\frac{1}{2}\right)}{\Gamma\left(m-\frac{p+1}{2}\right)\pi^{p/2}} \left[1+(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^{-m+\frac{1}{2}}$$
$$\times \int_{-\infty}^{\alpha^T(\mathbf{y}-\mu)} \frac{\Gamma(m)\left[1+(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^{m-\frac{1}{2}}}{\sqrt{\pi}\Gamma\left(m-\frac{1}{2}\right)\left[1+(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})+r^2\right]^m} \mathrm{d}r,$$

with

$$g^{(p+1)}(u) = \frac{\Gamma(m)}{\Gamma\left(m-\frac{p+1}{2}\right)\pi^{(p+1)/2}}(1+u)^{-m},\ m > \frac{p+1}{2},$$

then the first four moments of $\mathbf{Y}$ are expressed as follows:

(1) $M_1 = \boldsymbol{\mu} + \dfrac{\Gamma\left(m-1-\frac{p}{2}\right)}{\sqrt{\pi}\Gamma\left(m-\frac{p+1}{2}\right)}\boldsymbol{\delta},\ m > \dfrac{p}{2}+1;$

(2) $M_2 = \boldsymbol{\mu}\boldsymbol{\mu}^T + \dfrac{\Gamma\left(m-1-\frac{p}{2}\right)}{\sqrt{\pi}\Gamma\left(m-\frac{p+1}{2}\right)}\left(\boldsymbol{\mu}\boldsymbol{\delta}^T+\boldsymbol{\delta}\boldsymbol{\mu}^T\right) + \dfrac{1}{2m-p-3}\boldsymbol{\Omega},\ m > \dfrac{p+3}{2};$

(3)
$$\begin{aligned} M_3 = {} & \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu} + \frac{\Gamma\left(m-1-\frac{p}{2}\right)}{\sqrt{\pi}\Gamma\left(m-\frac{p+1}{2}\right)}[\boldsymbol{\delta}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}+\boldsymbol{\mu}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\mu} \\ & +\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\delta}] + \frac{1}{2m-p-3}\left[\boldsymbol{\Omega}\otimes\boldsymbol{\mu}+\boldsymbol{\mu}\otimes\boldsymbol{\Omega}+vec(\boldsymbol{\Omega})\otimes\boldsymbol{\mu}^T\right] \\ & + \frac{\Gamma\left(m-2-\frac{p}{2}\right)}{2^{\frac{3}{2}}\Gamma\left(m-\frac{p+1}{2}\right)}\sqrt{\frac{2}{\pi}}[\boldsymbol{\delta}\otimes\boldsymbol{\Omega}+vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T+\left(\mathbf{I}_p\otimes\boldsymbol{\delta}\right)\boldsymbol{\Omega} \\ & -\left(\mathbf{I}_p\otimes\boldsymbol{\delta}\right)\left(\boldsymbol{\delta}\otimes\boldsymbol{\delta}^T\right)],\ m > \frac{p}{2}+2; \end{aligned}$$

(4)
$$\begin{aligned} M_4 = {} & \boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T + \frac{\Gamma\left(m-1-\frac{p}{2}\right)}{\sqrt{\pi}\Gamma\left(m-\frac{p+1}{2}\right)}G_1 + \frac{1}{2m-p-3}G_2 \\ & + \frac{\Gamma\left(m-2-\frac{p}{2}\right)}{2^{\frac{3}{2}}\Gamma\left(m-\frac{p+1}{2}\right)}\sqrt{\frac{2}{\pi}}G_3 + \frac{1}{(2m-p-3)(2m-p-5)}G_4,\ m > \frac{p+5}{2}. \end{aligned}$$

Example 3.3 If $\mathbf{Y}=\{\mathbf{X}|X_0>0\}\sim SPII\left(\boldsymbol{\mu},\boldsymbol{\Omega},\boldsymbol{\delta},g^{(p+1)}\right)$ and its joint density function is given by

$$f_{\mathbf{Y}}(\mathbf{y})=2|\boldsymbol{\Omega}|^{-\frac{1}{2}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+\frac{3}{2}\right)\pi^{p/2}}s^{-m+\frac{p+1}{2}}\left[s-(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^{m+\frac{1}{2}}$$
$$\times\int_{-\sqrt{s}}^{\alpha^T(\mathbf{y}-\mu)}\frac{\Gamma\left(m+\frac{3}{2}\right)\left[s-r^2-(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^m}{\sqrt{\pi}\Gamma(m+1)\left[s-(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]^{m+\frac{1}{2}}}\mathrm{d}r,$$

with

$$g^{(p+1)}(u)=\frac{\Gamma\left[m+1+\frac{p+1}{2}\right]}{\Gamma(m+1)\pi^{(p+1)/2}}s^{-\left(m+\frac{p+1}{2}\right)}(s-u)^m,\ 0<u<s,\ m>-1,\ s>0,$$

then the first four moments of $\mathbf{Y}$ are expressed as follows:

(1) $M_1=\boldsymbol{\mu}+\sqrt{\frac{s}{\pi}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+2+\frac{p}{2}\right)}\boldsymbol{\delta};$

(2) $M_2=\boldsymbol{\mu}\boldsymbol{\mu}^T+\sqrt{\frac{s}{\pi}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+2+\frac{p}{2}\right)}\left(\boldsymbol{\mu}\boldsymbol{\delta}^T+\boldsymbol{\delta}\boldsymbol{\mu}^T\right)+\frac{s}{2m+p+3}\boldsymbol{\Omega};$

(3)
$$M_3=\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}+\sqrt{\frac{s}{\pi}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+2+\frac{p}{2}\right)}[\boldsymbol{\delta}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}+\boldsymbol{\mu}\otimes\boldsymbol{\delta}^T\otimes\boldsymbol{\mu}$$
$$+\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\delta}]+\frac{s}{2m+p+3}\left[\boldsymbol{\Omega}\otimes\boldsymbol{\mu}+\boldsymbol{\mu}\otimes\boldsymbol{\Omega}+vec(\boldsymbol{\Omega})\otimes\boldsymbol{\mu}^T\right]$$
$$+\left(\frac{s}{2}\right)^{\frac{3}{2}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+3+\frac{p}{2}\right)}\sqrt{\frac{2}{\pi}}[\boldsymbol{\delta}\otimes\boldsymbol{\Omega}+vec(\boldsymbol{\Omega})\boldsymbol{\delta}^T+\left(\mathbf{I}_p\otimes\boldsymbol{\delta}\right)\boldsymbol{\Omega}$$
$$-\left(\mathbf{I}_p\otimes\boldsymbol{\delta}\right)\left(\boldsymbol{\delta}\otimes\boldsymbol{\delta}^T\right)];$$

(4)
$$M_4=\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T+\sqrt{\frac{s}{\pi}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+2+\frac{p}{2}\right)}G_1+\frac{s}{2m+p+3}G_2$$
$$+\left(\frac{s}{2}\right)^{\frac{3}{2}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+3+\frac{p}{2}\right)}\sqrt{\frac{2}{\pi}}G_3+\frac{s^2}{(2m+p+3)(2m+p+5)}G_4.$$

4. Song's measure in the skew-elliptical distributions

Based on the λ order entropy, Song [8] has defined a general measure of the distribution shape known as Song's measure:

$$S(f) = -2\frac{\mathrm{d}}{\mathrm{d}\lambda}J_R(1) = Var[\log(f(X))], \tag{4.1}$$

where f denotes a univariate or multivariate density function, $J_R(\lambda) = \frac{1}{1-\lambda}\log G(\lambda)$, $\lambda > 0$, and $\lambda \neq 1$, $G(\lambda) = \int f^{\lambda}\mathrm{d}x$.

Zografos [1] calculated Song's kurtosis measures of multivariate t, multivariate Pearson VII and multivariate Pearson II distributions respectively, and the corresponding specific results are as follows:
(1) Multivariate t-distribution:

$$S(f) = \frac{(v+p)^2}{4}\left\{\Psi'\left(\frac{v}{2}\right) - \Psi'\left(\frac{v+p}{2}\right)\right\};$$

(2) Multivariate Pearson type VII distribution:

$$S(f) = m^2\left\{\Psi'\left(m-\frac{p}{2}\right) - \Psi'(m)\right\},\ m > \frac{p}{2};$$

(3) Multivariate Pearson type II distribution:

$$S(f) = m^2\left\{\Psi'(m+1) - \Psi'\left(\frac{p}{2}+m+1\right)\right\},\ m > -1,$$

where $\Psi(z) = (\mathrm{d}/\mathrm{d}z)\log\Gamma(z),\ \Psi'(z) = \left(\mathrm{d}^2/\mathrm{d}z^2\right)\log\Gamma(z)$.

Since finding the exact value of Song's measure of skew-elliptical distribution is very complicated and difficult to obtain, it is considered to approximate the variance of the random variable function by using delta method and Taylor expansion (similar to that used in Balakrishnan and Scarpa [14]).

The Taylor expansion of $h(x)$ at $x = a$ is

$$h(x) = h(a) + h'(a)(x-a) + \frac{h''(a)}{2!}(x-a)^2 + o\left[(x-a)^2\right],$$

and the approximate formula is

$$h(x) \approx h(a) + h'(a)(x-a).$$

If $E(X) = \mu,\ Var(X) = \sigma^2$, then

$$h(X) \approx h(\mu) + h'(\mu)(X-\mu).$$

According to the property of the variance, we get

$$Var[h(X)] \approx Var\left[h(\mu) + h'(\mu)(X-\mu)\right] = [h'(\mu)]^2\sigma^2.$$

If $\mathbf{X}$ is a p-dimensional random vector, $E(\mathbf{X}) = \boldsymbol{\mu}$ and $Var(\mathbf{X}) = \boldsymbol{\Sigma}$, then

$$Var[h(\mathbf{X})] \approx [h'(\boldsymbol{\mu})]^T\boldsymbol{\Sigma}h'(\boldsymbol{\mu}). \tag{4.2}$$

Proposition 4.1. *If* $\mathbf{Y} \sim SE_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$ *and has the density function form of (3.1), then Song's kurtosis measure is*

$$S(f) \approx \left\{\frac{2\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\hat{g}'^{(p)}\left[(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{\hat{g}^{(p)}\left[(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]} + \frac{f_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{F_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\}^T$$
$$Var(\mathbf{Y})\left\{\frac{2\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\hat{g}'^{(p)}\left[(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{\hat{g}^{(p)}\left[(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]} + \frac{f_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{F_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\}, \tag{4.3}$$

where $\boldsymbol{\mu}_{\mathbf{Y}}$ is the mean vector of $\mathbf{Y}$.

Proof. From (4.1) and (4.2), let $H(\mathbf{y}) = \log f(\mathbf{y})$, we obtain

$$S(f) = Var\left[\log f(\mathbf{Y})\right] = Var\left[H(\mathbf{Y})\right] \approx \left[H'\left(E(\mathbf{Y})\right)\right]^T Var(\mathbf{Y})\left[H'\left(E(\mathbf{Y})\right)\right], \tag{4.4}$$

$$\begin{aligned} H(\mathbf{y}) &= \log\left(2f_{\hat{g}^{(p)}}(\mathbf{y})F_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right]\right) \\ &= \log 2 + \log|\boldsymbol{\Omega}|^{\frac{1}{2}}\hat{g}^{(p)}\left[(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right] + \log F_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right] \\ &= \log 2 - \frac{1}{2}\log|\boldsymbol{\Omega}| + \log\hat{g}^{(p)}\left[(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right] + \log F_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right]. \end{aligned} \tag{4.5}$$

Because $F_{\tilde{g}_{q(\mathbf{y})}}$ is the distribution function of $EC_1\left(0, 1, \tilde{g}_{q(\mathbf{y})}\right)$,

$$\begin{aligned} f_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right] &= \tilde{g}_{q(\mathbf{y})}\left[(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\alpha}\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right] \\ &= \frac{g^{(p+1)}\left[(\mathbf{y}-\boldsymbol{\mu})^T\left(\boldsymbol{\alpha}\boldsymbol{\alpha}^T+\boldsymbol{\Omega}^{-1}\right)(\mathbf{y}-\boldsymbol{\mu})\right]}{\hat{g}^{(p)}\left[(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]}. \end{aligned} \tag{4.6}$$

From (4.5) and (4.6), we obtain

$$H'(\mathbf{y}) = \frac{2\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\hat{g}'^{(p)}\left[(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]}{\hat{g}^{(p)}\left[(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right]} + \frac{f_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right]}{F_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\mathbf{y}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha},$$

$$\begin{aligned} H'(E(\mathbf{Y})) = &\frac{2\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\hat{g}'^{(p)}\left[(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{\hat{g}^{(p)}\left[(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]} \\ &+ \frac{f_{\tilde{g}_{q(\mathbf{y})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{F_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}. \end{aligned}$$

Using $H'\left(E(\mathbf{Y})\right)$ and (4.4), complete the proof of Proposition 4.1. □

Example 4.1 If $\mathbf{Y} \sim St_p(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)})$, then

$$\begin{aligned} S(f) \approx &\left\{-\frac{(\nu+p)\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})}{\nu+(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})} + \frac{t^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{T^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\}^T \\ &\times Var(\mathbf{Y})\left\{-\frac{(\nu+p)\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_Y-\boldsymbol{\mu})}{\nu+(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})} + \frac{t^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{T^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\}, \end{aligned}$$

where

$$\boldsymbol{\mu}_{\mathbf{Y}} - \boldsymbol{\mu} = \sqrt{\frac{\nu}{\pi}}\frac{\Gamma[(\nu-1)/2]}{\Gamma(\nu/2)}\boldsymbol{\delta},\ \nu > 1,$$

$$Var(\mathbf{Y}) = \frac{\nu}{\nu-2}\boldsymbol{\Omega} - [\sqrt{\frac{\nu}{\pi}}\frac{\Gamma[(\nu-1)/2]}{\Gamma(\nu/2)}]^2\boldsymbol{\delta\delta}^T,\ \nu > 2,$$

$$t^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right] = \frac{\Gamma\left(\frac{\nu+p+1}{2}\right)\left(1+\frac{(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})}{\nu}\right)^{\frac{\nu+p}{2}}}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu+p}{2}\right)\left(1+\frac{(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T(\alpha\alpha^T+\boldsymbol{\Omega}^{-1})(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})}{\nu}\right)^{(\nu+p+1)/2}},$$

$$T^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right] = \int_{-\infty}^{\alpha^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})} t^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}(r)\mathrm{d}r.$$

Proof. It is proved by using Example 3.1 and Proposition 4.1. □

Example 4.2 If $\mathbf{Y} \sim SPVII_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, then

$$S(f) \approx \left\{-\frac{(2m-1)\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})}{1+(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})} + \frac{pVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{PVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\}^T$$
$$\times Var(\mathbf{Y})\left\{-\frac{(2m-1)\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})}{1+(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})} + \frac{pVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}{PVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\},$$

where

$$\boldsymbol{\mu}_{\mathbf{Y}} - \boldsymbol{\mu} = \frac{\Gamma[(m-1)-p/2]}{\sqrt{\pi}\Gamma(m-(p+1)/2)}\boldsymbol{\delta},\ m > \frac{p}{2}+1,$$

$$Var(\mathbf{Y}) = \frac{1}{2m-p-3}\boldsymbol{\Omega} - \left[\frac{\Gamma[(m-1)-p/2]}{\sqrt{\pi}\Gamma(m-(p+1)/2)}\right]^2\boldsymbol{\delta\delta}^T,\ m > \frac{p+3}{2},$$

$$pVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right] = \frac{\Gamma(m)\left[1+(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]^{m-\frac{1}{2}}}{\Gamma(m-\frac{1}{2})\sqrt{\pi}\left[1+(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})^T(\alpha\alpha^T+\boldsymbol{\Omega}^{-1})(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right]^m},$$

$$PVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})\right] = \int_{-\infty}^{\alpha^T(\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\mu})} pVII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}(r)\mathrm{d}r.$$

Example 4.3 If $\mathbf{Y} \sim SPII_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \delta, g^{(p+1)}\right)$, then

$$S(f) \approx \left\{-\frac{(2m+1)\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_Y-\boldsymbol{\mu})}{s-(\boldsymbol{\mu}_Y-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_Y-\boldsymbol{\mu})} + \frac{pII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y-\boldsymbol{\mu})\right]}{PII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\}^T$$
$$\times Var(\mathbf{Y})\left\{-\frac{(2m+1)\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_Y-\boldsymbol{\mu})}{s-(\boldsymbol{\mu}_Y-\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_Y-\boldsymbol{\mu})} + \frac{pII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y-\boldsymbol{\mu})\right]}{PII^*_{\tilde{g}_{q(\boldsymbol{\mu}_{\mathbf{Y}})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y-\boldsymbol{\mu})\right]}\boldsymbol{\alpha}\right\},$$

where

$$\boldsymbol{\mu}_Y - \boldsymbol{\mu} = \sqrt{\frac{s}{\pi}}\frac{\Gamma[m+1+(p+1)/2]}{\Gamma(m+2+p/2)}\boldsymbol{\delta},\ m > -1,\ s > 0,$$

$$Var(\mathbf{Y}) = \frac{s}{2m+p+3}\boldsymbol{\Omega} - \left[\sqrt{\frac{s}{\pi}}\frac{\Gamma[m+1+(p+1)/2]}{\Gamma(m+2+p/2)}\right]^2 \boldsymbol{\delta}\boldsymbol{\delta}^T,\ m > -1,\ s > 0,$$

$$pII^*_{\tilde{g}_{q(\boldsymbol{\mu}_\mathbf{Y})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y - \boldsymbol{\mu})\right] = \frac{\Gamma\left(m+\frac{3}{2}\right)\left[s - ((\boldsymbol{\mu}_Y - \boldsymbol{\mu})^T(\boldsymbol{\alpha}\boldsymbol{\alpha}^T + \boldsymbol{\Omega}^{-1})(\boldsymbol{\mu}_Y - \boldsymbol{\mu}))\right]^m}{\Gamma(m+1)\sqrt{\pi}\left[s - (\boldsymbol{\mu}_Y - \boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}_Y - \boldsymbol{\mu})\right]^{m+\frac{1}{2}}},$$

$$PII^*_{\tilde{g}_{q(\boldsymbol{\mu}_\mathbf{Y})}}\left[\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y - \boldsymbol{\mu})\right] = \int_{-\sqrt{s}}^{\boldsymbol{\alpha}^T(\boldsymbol{\mu}_Y - \boldsymbol{\mu})} pII^*_{\tilde{g}_{q(\boldsymbol{\mu}_\mathbf{Y})}}(r)\mathrm{d}r.$$

5. Mardia's measure of kurtosis in skew-elliptical distributions

In Section 2, we have considered Marida's kurtosis measure:

$$\beta_{2,p} = E\left[(\mathbf{X} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})\right]^2$$

and the result for elliptical distribution has been given by Zografos [1]. Next, we consider how to calculate Mardia's kurtosis measure for skew-elliptical distributions.

Lemma 5.1. *If* $\mathbf{Y} \sim SE_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$*, then*

$$\tilde{\mathbf{Y}} = \boldsymbol{\Sigma}_\mathbf{Y}^{-\frac{1}{2}}(\mathbf{Y} - \boldsymbol{\mu}_\mathbf{Y}) \sim SE_p\left(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Omega}}, \tilde{\boldsymbol{\delta}}, g^{(p+1)}\right),$$

where $\boldsymbol{\mu}_\mathbf{Y}$ *is the mean vector of* $\mathbf{Y}$*,* $\boldsymbol{\Sigma}_\mathbf{Y}$ *is the covariance matrix of* $\mathbf{Y}$*,* $\tilde{\boldsymbol{\mu}} = \boldsymbol{\Sigma}_\mathbf{Y}^{-\frac{1}{2}}(\boldsymbol{\mu} - \boldsymbol{\mu}_\mathbf{Y})$*,* $\tilde{\boldsymbol{\Omega}} = \boldsymbol{\Sigma}_Y^{-\frac{1}{2}}\boldsymbol{\Omega}\left(\boldsymbol{\Sigma}_Y^{-\frac{1}{2}}\right)^T$*,* $\tilde{\boldsymbol{\delta}} = \boldsymbol{\Sigma}_\mathbf{Y}^{-\frac{1}{2}}\boldsymbol{\delta}$*.*

Lemma 5.1 is a special case of Branco and Dey [25] for $\boldsymbol{C} = \boldsymbol{\Sigma}_Y^{-1/2}$ and $\boldsymbol{b} = \boldsymbol{\Sigma}_Y^{-1/2}\boldsymbol{\mu}_Y$. From the Lemma 5.1, it can be known that the Mardia's kurtosis measure of skew-elliptical distribution

$$\beta_{2,p} = E\left[(\boldsymbol{Y} - \boldsymbol{\mu}_Y)^T\boldsymbol{\Sigma}_\mathbf{Y}{}^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}_Y)\right]^2 = E\left(\tilde{\mathbf{Y}}^T\tilde{\mathbf{Y}}\right)^2$$

turns out to be the expectation of the square of a quadratic form of a skew-elliptical distribution, that is,

$$\beta_{2,p} = E\left(\tilde{\mathbf{Y}}^T\mathbf{I}_p\tilde{\mathbf{Y}}\right)^2,$$

where $\mathbf{I}_p$ is the p-dimensional unit matrix.

Lemma 5.2. *(Schott [32]) Let* $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ *be matrices of sizes* $m \times n,\ p \times q,\ n \times p,\ q \times m$*,* $\mathbf{U}_{nn}$ *be the* $n^2 \times n^2$ *commutation matrix,* $\mathbf{U}_{pm}$ *be the* $pm \times pm$ *commutation matrix,* $\mathbf{U}_{mp} = \mathbf{U}_{pm}^T$*,* $\mathbf{x}$ *be an* $m \times 1$ *vector,* $\mathbf{y}$ *be a* $p \times 1$ *vector. Then*

$$tr(\mathbf{A} \otimes \mathbf{B}) = tr(\mathbf{A})tr(\mathbf{B});$$

$$tr\left((\mathbf{A} \otimes \mathbf{B})\mathbf{U}_{mn}\right) = tr(\mathbf{AB}),\ if\ p = n\ and\ q = m;$$

$$\mathbf{U}_{pm}(\mathbf{A} \otimes \mathbf{y}) = \mathbf{y} \otimes \mathbf{A};\ \ \mathbf{U}_{mp}(\mathbf{y} \otimes \mathbf{A}) = \mathbf{A} \otimes \mathbf{y};$$

$$\mathbf{U}_{pm}(\mathbf{x} \otimes \mathbf{y}) = \mathbf{y} \otimes \mathbf{x};\ \ tr\left(\mathbf{ACBD}\right) = \left(vec(\mathbf{A}^T)\right)^T\left(\mathbf{D}^T \otimes \mathbf{C}\right)vec(\mathbf{B}).$$

Proposition 5.1. *If* $\mathbf{Y} \sim SE_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$ *and* $\mathbf{A}, \mathbf{B}$ *are the* $p \times p$ *dimensional symmetric matrix, then*

$$\begin{aligned}
E\left(\mathbf{Y}^T\mathbf{A}\mathbf{Y}\right)\left(\mathbf{Y}^T\mathbf{B}\mathbf{Y}\right) = {} & \left(\boldsymbol{\mu}^T\boldsymbol{A\mu}\right)\left(\boldsymbol{\mu}^T\boldsymbol{B\mu}\right) + \sqrt{\frac{2}{\pi}}\frac{E(R)}{E(R_0)}[2\left(\boldsymbol{\mu}^T\boldsymbol{A\delta}\right)\left(\boldsymbol{\mu}^T\boldsymbol{B\mu}\right) \\
& + 2\left(\boldsymbol{\mu}^T\boldsymbol{A\mu}\right)\left(\boldsymbol{\mu}^T\boldsymbol{A\delta}\right)] + \frac{E\left(R^2\right)}{E\left(R_0^2\right)}[\left(\boldsymbol{\mu}^T\boldsymbol{B\mu}\right)tr(\mathbf{A}\boldsymbol{\Omega}) \\
& + 2\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\Omega}\boldsymbol{B\mu} + 2\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\Omega}\mathbf{A}\boldsymbol{\mu} + \left(\boldsymbol{\mu}^T\boldsymbol{A\mu}\right)tr(\mathbf{B}\boldsymbol{\Omega})] \\
& + \sqrt{\frac{2}{\pi}}\frac{E\left(R^3\right)}{E\left(R_0^3\right)}[4\boldsymbol{\mu}^T\boldsymbol{A\Omega B\delta} + 4\boldsymbol{\mu}^T\boldsymbol{B\Omega A\delta} \\
& + 2\boldsymbol{\mu}^T\boldsymbol{B\delta}tr(\mathbf{A}\boldsymbol{\Omega}) + 2\boldsymbol{\mu}^T\boldsymbol{A\delta}tr(\mathbf{B}\boldsymbol{\Omega}) - 2\left(\boldsymbol{\delta}^T\boldsymbol{A\delta}\right) \\
& \times\left(\boldsymbol{\mu}^T\boldsymbol{B\delta}\right) - 2\left(\boldsymbol{\mu}^T\boldsymbol{A\delta}\right)\left(\boldsymbol{\delta}^T\boldsymbol{B\delta}\right)] + \frac{E\left(R^4\right)}{E\left(R_0^4\right)}[tr(\mathbf{A}\boldsymbol{\Omega}) \\
& \times tr(\mathbf{B}\boldsymbol{\Omega}) + 2tr(\mathbf{A}\boldsymbol{\Omega}\mathbf{B}\boldsymbol{\Omega})].
\end{aligned}$$

Proof. From $E\left(\mathbf{Y}^T\mathbf{A}\mathbf{Y}\right)\left(\mathbf{Y}^T\mathbf{B}\mathbf{Y}\right) = tr\left[(\mathbf{A}\otimes\mathbf{B})M_4\right]$ (See Genton et al. [12]) and Theorem 3.1 and Lemma 5.2, the following results can be obtained

$$\begin{aligned}
& tr\left[(\mathbf{A}\otimes\mathbf{B})\left(\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\right)\right] = tr\left(\boldsymbol{A\mu\mu}^T\otimes\boldsymbol{B\mu\mu}^T\right) = \left(\boldsymbol{\mu}^T\boldsymbol{A\mu}\right)\left(\boldsymbol{\mu}^T\boldsymbol{B\mu}\right); \\
& tr\left[(\mathbf{A}\otimes\mathbf{B})G_1\right] = 2\left(\boldsymbol{\mu}^T\boldsymbol{A\delta}\right)\left(\boldsymbol{\mu}^T\boldsymbol{B\mu}\right) + 2\left(\boldsymbol{\mu}^T\boldsymbol{A\mu}\right)\left(\boldsymbol{\mu}^T\boldsymbol{A\delta}\right); \\
& tr\left[(\mathbf{A}\otimes\mathbf{B})G_2\right] = \left(\boldsymbol{\mu}^T\boldsymbol{B\mu}\right)tr(\boldsymbol{A\Omega}) + 2\boldsymbol{\mu}^T\boldsymbol{A\Omega B\mu} + 2\boldsymbol{\mu}^T\boldsymbol{B\Omega A\mu} \\
& \qquad + \left(\boldsymbol{\mu}^T\boldsymbol{A\mu}\right)tr(\boldsymbol{B\Omega}); \\
& tr[(\mathbf{A}\otimes\mathbf{B})G_3] = 4\boldsymbol{\mu}^T\boldsymbol{A\Omega B\delta} + 4\boldsymbol{\mu}^T\boldsymbol{B\Omega A\delta} + 2\boldsymbol{\mu}^T\boldsymbol{B\delta}tr(\boldsymbol{A\Omega}) \\
& \qquad + 2\boldsymbol{\mu}^T\boldsymbol{A\delta}tr(\mathbf{B}\boldsymbol{\Omega}) - 2\left(\boldsymbol{\delta}^T\boldsymbol{A\delta}\right)\left(\boldsymbol{\mu}^T\boldsymbol{B\delta}\right) - 2\left(\boldsymbol{\mu}^T\boldsymbol{A\delta}\right)\left(\boldsymbol{\delta}^T\boldsymbol{B\delta}\right); \\
& tr[(\mathbf{A}\otimes\mathbf{B})G_4] = tr(\mathbf{A}\boldsymbol{\Omega})tr(\mathbf{B}\boldsymbol{\Omega}) + 2tr(\mathbf{A}\boldsymbol{\Omega}\mathbf{B}\boldsymbol{\Omega}); \\
& E\left(\mathbf{Y}^T\mathbf{A}\mathbf{Y}\right)\left(\mathbf{Y}^T\mathbf{B}\mathbf{Y}\right) = tr[(\mathbf{A}\otimes\mathbf{B})M_4] = tr\left[(\mathbf{A}\otimes\mathbf{B})\left(\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\otimes\boldsymbol{\mu}\otimes\boldsymbol{\mu}^T\right)\right] \\
& \qquad + \sqrt{\frac{2}{\pi}}\frac{E(R)}{E(R_0)}tr[(\mathbf{A}\otimes\mathbf{B})G_1] + \frac{E\left(R^2\right)}{E\left(R_0^2\right)}tr[(\mathbf{A}\otimes\mathbf{B})G_2] \\
& \qquad + \sqrt{\frac{2}{\pi}}\frac{E\left(R^3\right)}{E\left(R_0^3\right)}tr[(\mathbf{A}\otimes\mathbf{B})G_3] + \frac{E\left(R^4\right)}{E\left(R_0^4\right)}tr[(\mathbf{A}\otimes\mathbf{B})G_4].
\end{aligned}$$

□

Corollary 1. *If* $\tilde{\mathbf{Y}} \sim SE_p\left(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Omega}}, \tilde{\boldsymbol{\delta}}, g^{(p+1)}\right)$*, then*

$$E\left(\tilde{\mathbf{Y}}^T\tilde{\mathbf{Y}}\right)^2 = \left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)^2 + 4\sqrt{\frac{2}{\pi}}\frac{E(R)}{E(R_0)}\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right) + 2\frac{E\left(R^2\right)}{E\left(R_0^2\right)}[\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)tr\left(\tilde{\boldsymbol{\Omega}}\right)$$
$$+ 2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\mu}}] + 4\sqrt{\frac{2}{\pi}}\frac{E\left(R^3\right)}{E\left(R_0^3\right)}[2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}tr\left(\tilde{\boldsymbol{\Omega}}\right) - \left(\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)]$$
$$+ \frac{E\left(R^4\right)}{E\left(R_0^4\right)}\left[\left(tr\left(\tilde{\boldsymbol{\Omega}}\right)\right)^2 + 2tr\left(\tilde{\boldsymbol{\Omega}}^2\right)\right],$$

where $\tilde{\boldsymbol{\mu}}$, $\tilde{\boldsymbol{\Omega}}$, $\tilde{\boldsymbol{\delta}}$ *is consistent with the form in Lemma 5.1.*

Proof. By Proposition 5.1, let $\mathbf{A} = \mathbf{B} = \mathbf{I}_p$, the conclusion can be proved. □

From this, we can get Mardia's kurtosis measures of skew-elliptical distributions. Next, we apply Corollary 1 and Proposition 5.1 to give three examples:

Example 5.1 If $\mathbf{Y} \sim St_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, then Mardia's kurtosis measure of $\mathbf{Y}$ is

$$\beta_{2,p} = \left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)^2 + 4\sqrt{\frac{\nu}{\pi}}\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right) + \frac{2\nu}{\nu-2}\left[\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)tr\left(\tilde{\boldsymbol{\Omega}}\right) + 2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\mu}}\right]$$
$$+ \frac{2\nu^{\frac{3}{2}}}{\sqrt{\pi}}\frac{\Gamma\left(\frac{\nu-1}{2}-1\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left[2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}tr\left(\tilde{\boldsymbol{\Omega}}\right) - \left(\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\right]$$
$$+ \frac{\nu^2}{(\nu-2)(\nu-4)}\left[\left(tr\left(\tilde{\boldsymbol{\Omega}}\right)\right)^2 + 2tr\left(\tilde{\boldsymbol{\Omega}}^2\right)\right],\ \nu > 4,$$

where

$$\tilde{\boldsymbol{\mu}} = -\sqrt{\frac{\nu}{\pi}}\frac{\Gamma[(\nu-1)/2]}{\Gamma(\nu/2)}\left\{\frac{\nu}{\nu-2}\boldsymbol{\Omega} - \left[\sqrt{\frac{\nu}{\pi}}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\right]^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right\}^{-\frac{1}{2}}\boldsymbol{\delta};$$

$$\tilde{\boldsymbol{\Omega}} = \left\{\frac{\nu}{\nu-2}\boldsymbol{\Omega} - \left[\sqrt{\frac{\nu}{\pi}}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\right]^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right\}^{-\frac{1}{2}}\boldsymbol{\Omega}$$
$$\cdot\left\{\left(\frac{\nu}{\nu-2}\boldsymbol{\Omega} - \left[\sqrt{\frac{\nu}{\pi}}\frac{\Gamma[(\nu-1)/2]}{\Gamma(\nu/2)}\right]^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right)^{-\frac{1}{2}}\right\}^T;$$

$$\tilde{\boldsymbol{\delta}} = \left\{\frac{\nu}{\nu-2}\boldsymbol{\Omega} - \left[\sqrt{\frac{\nu}{\pi}}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\right]^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right\}^{-\frac{1}{2}}\boldsymbol{\delta}.$$

Example 5.2 If $\mathbf{Y} \sim SPVII_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, then Mardia's kurtosis measure of $\mathbf{Y}$ is

$$\begin{aligned}\beta_{2,p} =& \left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)^2 + 4c_1\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right) + \frac{2}{2m-p-3}\left[\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)tr\left(\tilde{\boldsymbol{\Omega}}\right) + 2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\mu}}\right]\\ &+ \frac{2\Gamma\left(m-2-\frac{p}{2}\right)}{\sqrt{\pi}\Gamma\left(m-\frac{p+1}{2}\right)}\left[2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}tr\left(\tilde{\boldsymbol{\Omega}}\right) - \left(\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\right]\\ &+ \frac{1}{(2m-p-3)(2m-p-5)}\left[\left(tr\left(\tilde{\boldsymbol{\Omega}}\right)\right)^2 + 2tr\left(\tilde{\boldsymbol{\Omega}}^2\right)\right],\ m > \frac{p+5}{2},\end{aligned}$$

where

$$\begin{aligned}c_1 =& \frac{\Gamma\left(m-1-\frac{p}{2}\right)}{\sqrt{\pi}\Gamma\left(m-\frac{p+1}{2}\right)},\\ \tilde{\boldsymbol{\mu}} =& -c_1\left[\frac{1}{2m-p-3}\boldsymbol{\Omega} - c_1^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right]^{-\frac{1}{2}}\boldsymbol{\delta},\\ \tilde{\boldsymbol{\Omega}} =& \left[\frac{1}{2m-p-3}\boldsymbol{\Omega} - c_1^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right]^{-\frac{1}{2}}\boldsymbol{\Omega}\left[\left(\frac{1}{2m-p-3}\boldsymbol{\Omega} - c_1^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right)^{-\frac{1}{2}}\right]^T,\\ \tilde{\boldsymbol{\delta}} =& \left[\frac{1}{2m-p-3}\boldsymbol{\Omega} - c_1^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right]^{-\frac{1}{2}}\boldsymbol{\delta}.\end{aligned}$$

Example 5.3 If $\mathbf{Y} \sim SPII_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, then Mardia's kurtosis measure of $\mathbf{Y}$ is

$$\begin{aligned}\beta_{2,p} =& \left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)^2 + 4c_2\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right) + \frac{2s}{2m+p+3}\left[\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)tr\left(\tilde{\boldsymbol{\Omega}}\right) + 2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\mu}}\right]\\ &+ \frac{2s^{\frac{3}{2}}\Gamma\left(m+1+\frac{p+1}{2}\right)}{\sqrt{\pi}\Gamma\left(m+3+\frac{p}{2}\right)}\left[2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}tr\left(\tilde{\boldsymbol{\Omega}}\right) - \left(\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\right]\\ &+ \frac{s^2}{(2m+p+3)(2m+p+5)}\left[\left(tr(\tilde{\boldsymbol{\Omega}})\right)^2 + 2tr(\tilde{\boldsymbol{\Omega}}^2)\right],\ m > -1,\ s > 0,\end{aligned}$$

where

$$\begin{aligned}c_2 =& \sqrt{\frac{s}{\pi}}\frac{\Gamma\left(m+1+\frac{p+1}{2}\right)}{\Gamma\left(m+2+\frac{p}{2}\right)},\\ \tilde{\boldsymbol{\mu}} =& -c_2\left[\frac{s}{2m+p+3}\boldsymbol{\Omega} - c_2^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right]^{-\frac{1}{2}}\boldsymbol{\delta},\\ \tilde{\boldsymbol{\Omega}} =& \left[\frac{s}{2m+p+3}\boldsymbol{\Omega} - c_2^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right]^{-\frac{1}{2}}\boldsymbol{\Omega}\left[\left(\frac{s}{2m+p+3}\boldsymbol{\Omega} - c_2^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right)^{-\frac{1}{2}}\right]^T,\\ \tilde{\boldsymbol{\delta}} =& \left[\frac{s}{2m+p+3}\boldsymbol{\Omega} - c_2^2\boldsymbol{\delta}\boldsymbol{\delta}^T\right]^{-\frac{1}{2}}\boldsymbol{\delta}.\end{aligned}$$

6. Numerical analysis

Next, we consider the comparison of Mardia's and Song's measures of above distributions at different parameters values. First we have known that $\beta_{2,p}$ and Song's measure do not always exist. For example, when $\nu \leq 4$, $\beta_{2,p}$ of skew-t distribution does not exist. Therefore, we let the parameters of skew-t distribution: ν=10, 20, 30, 50, 100, the parameters of skew-Pearson type VII distribution: m=10, 20, 30, 50 and the parameters of skew-Pearson type II distribution: s=3, m=2, 4, 6, 12. We choose p=3, 6, 12, and the results are given in Table 2 and Table 3.

Table 2. Comparison of kurtosis measures at different dimensions and ν-values of skew-t distribution.

ST / p	ν	$\mathcal{S}(f)$	$\beta_{2,p}$
p=3	ν=10	0.0357	20.1682
	ν=20	0.0100	16.9255
	ν=30	0.0055	16.1882
	ν=50	0.0031	15.6782
	ν=100	0.0018	15.3343
p=6	ν=10	0.1804	64.4543
	ν=20	0.0501	54.1396
	ν=30	0.0278	51.7881
	ν=50	0.0154	50.1601
	ν=100	0.0088	49.0618
p=12	ν=10	0.9520	225.1901
	ν=20	0.2651	189.3568
	ν=30	0.1441	181.1648
	ν=50	0.0773	175.4872
	ν=100	0.0423	171.6536

From Table 2, we can observe that Song's measures and $\beta_{2,p}$ of skew-t distribution decrease as the degrees of ν increase. This phenomenon is justified because the tails become lighter as the degree of freedom ν increases. Moreover, taking into account that tails become lighter and Song's measure tends to zero when ν increases, we may infer that Song's measure is mainly concerned to the tails of a distribution.

Table 3 shows values of $\beta_{2,p}$ and Song's measures for some values of the parameter m and the dimensions p of skew-Pearson type VII and skew-Pearson type II distributions. By observing Tables 1 and 2, we can find that $\beta_{2,p}$ and Song's measures increase when the dimension p increases.

All the above phenomena suggest that Song's measure is more sensitive to the tails of a distribution while $\beta_{2,p}$ is more sensitive in the centre, which is consistent with that of Zografos [1] described. Both kurtosis measures provide useful information about the kurtosis of skew-elliptical distributions.

Table 3. Values of $\beta_{2,p}$ and Song's measures for skew-Pearson type VII and skew-Pearson type II distributions.

SPVII / p	m	$\mathcal{S}(f)$	$\beta_{2,p}$
p=3	m=10	0.0145	17.5683
	m=20	0.0044	15.9679
	m=30	0.0028	15.6019
	m=50	0.0018	15.3481
p=6	m=10	0.1065	58.9302
	m=20	0.0246	51.4002
	m=30	0.0145	50.0308
	m=50	0.0092	49.1411
p=12	m=10	2.0827	283.3144
	m=20	0.1672	182.8716
	m=30	0.0827	176.0014
	m=50	0.0468	172.2055
SPII			
p=3	m=2	0.0083	12.9026
	m=4	0.0034	14.9109
	m=6	0.0016	16.6496
	m=12	0.0002	21.2602
p=6	m=2	0.0163	40.52733
	m=4	0.0079	44.1548
	m=6	0.0038	47.2754
	m=12	0.0003	55.3368
p=12	m=2	0.0275	145.2465
	m=4	0.0161	151.2772
	m=6	0.0088	156.5489
	m=12	0.0008	170.0363

Similar to Balakishnan and Scarpa [14], we consider the sample version for each of the measures considered as test statistics for the hypothesis of normal against the skew-normal distribution (SN). By proceeding via simulation, a sensitivity index (p-value) for these kurtosis measures can be provided by enumerating the number of samples from multivariate normal distribution having each index of kurtosis not exceeding the theoretical value obtained for the SN. For Song's measure, sensitivity is obtained by considering the reverse rejection regions. By Proposition 4.1 and Corollary 1, Song's and

Mardia's measures of kurtosis in skew-normal distribution are as follows.

If $\tilde{\mathbf{Y}} \sim SN_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\delta}, g^{(p+1)}\right)$, then

$$S(f) \approx \left\{-\sqrt{\frac{2}{\pi}}\boldsymbol{\Omega}^{-1}\boldsymbol{\delta} + \frac{\phi\left(\boldsymbol{\alpha}^T\boldsymbol{\delta}\right)}{\Phi\left(\boldsymbol{\alpha}^T\boldsymbol{\delta}\right)}\boldsymbol{\alpha}\right\}^T \times\left[\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^T\right]\left\{-\sqrt{\frac{2}{\pi}}\boldsymbol{\Omega}^{-1}\boldsymbol{\delta} + \frac{\phi\left(\boldsymbol{\alpha}^T\boldsymbol{\delta}\right)}{\Phi\left(\boldsymbol{\alpha}^T\boldsymbol{\delta}\right)}\boldsymbol{\alpha}\right\},$$

where

$$\boldsymbol{\delta} = \frac{\boldsymbol{\Omega}\boldsymbol{\alpha}}{(1+\boldsymbol{\alpha}^T\boldsymbol{\Omega}\boldsymbol{\alpha})^{1/2}}.$$

$$\beta_{2,p} = \left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)^2 + 4\sqrt{\frac{2}{\pi}}\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right) + 2[\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\mu}}\right)tr\left(\tilde{\boldsymbol{\Omega}}\right) + 2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\mu}}] + 4\sqrt{\frac{2}{\pi}}[2\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{\delta}} + \tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}tr\left(\tilde{\boldsymbol{\Omega}}\right) - \left(\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}}\right)\left(\tilde{\boldsymbol{\mu}}^T\tilde{\boldsymbol{\delta}}\right)] + \left(tr\left(\tilde{\boldsymbol{\Omega}}\right)\right)^2 + 2tr\left(\tilde{\boldsymbol{\Omega}}^2\right),$$

where

$$\tilde{\boldsymbol{\mu}} = -\sqrt{\frac{2}{\pi}}\left\{\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^T\right\}^{-\frac{1}{2}}\boldsymbol{\delta};$$

$$\tilde{\boldsymbol{\Omega}} = \left\{\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^T\right\}^{-\frac{1}{2}}\boldsymbol{\Omega}\left\{\left(\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^T\right)^{-\frac{1}{2}}\right\}^T;$$

$$\tilde{\boldsymbol{\delta}} = \left\{\boldsymbol{\Omega} - \frac{2}{\pi}\boldsymbol{\delta}\boldsymbol{\delta}^T\right\}^{-\frac{1}{2}}\boldsymbol{\delta}.$$

For sensitivity analysis, we simulated 1000 samples of size 100 from the multivariate normal model with parameter settings as listed in Table 4. For each sample, we computed every empirical index of kurtosis and counted the proportion of samples for which the kurtosis index fell in the rejection region. From Table 4, we can find that the measures of kurtosis of Mardia performs very well in most of the cases considered in terms of indication of kurtosis as well as in terms of sensitivity, but the Song's measure performs relatively average. Therefore, Marida's kurtosis measure is the one to be recommended for practical use.

Table 4. Kurtosis measures for test for some bivariate skew-normal distributions and p-values for test for bivariate normality against skew-normal.

Parameters			$\beta_{2,p}$	$\mathcal{S}(f)$
$\boldsymbol{\Omega}$	$\begin{pmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{pmatrix}$	kurtosis value:	13.527	0.0149
$\boldsymbol{\alpha}$	$\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$	p-value:	0.999	0.480
$\boldsymbol{\Omega}$	$\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$	kurtosis value:	10.412	0.0093
$\boldsymbol{\alpha}$	$\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$	p-value:	0.992	0.457
$\boldsymbol{\Omega}$	$\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$	kurtosis value:	9.2497	0.0049
$\boldsymbol{\alpha}$	$\begin{pmatrix} 0 \\ 0.5 \end{pmatrix}$	p-value:	0.950	0.032

7. Concluding remarks

In this paper, the Mardia's and Song's kurtosis measures of elliptical distributions obtained by Zografos [1] have been further generalized. The kurtosis measures and first fourth-order moments of skew-elliptical distributions with a concise structure were obtained in a more understandable way. By comparing Mardia's and Song measures for specific members of the skew-elliptical distribution family, it was found that Song's measure should be mainly used to describe the movement of the probability mass from the shoulders into the tails and $\beta_{2,p}$ should be used mainly to describe the similar movement from the shoulders into the centre of the distribution. We should consider both together when studying the kurtosis of the distribution, which will better help us understand the shape of the distribution.

For Song's kurtosis measure, we have used the delta method to give the approximate expression, which is not an accurate expression. How to accurately calculate Song's kurtosis measure of skew-elliptical distribution as well as the generalized skew-elliptical distribution (see, for instance, Zuo and Yin [33]) is the goal of further research, since at this moment we are not able to give a formal proof of such a problem.

Acknowledgments

The authors would like to thank the two anonymous referees for their helpful comments and suggestions on an earlier version of this manuscript which led to this improved version. The research was supported by the National Natural Science Foundation of China (No. 12071251).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. K. Zografos, On Mardia's and Song's measures of kurtosis in elliptical distributions, *J. Multivariate Anal.*, **99** (2008), 858–879. https://doi.org/10.1016/j.jmva.2007.05.001
2. B. C. Arnold, R. A. Groeneveld, Measuring skewness with respect to the mode, *The American Statisticia*, **49** (1995), 34–38. https://doi.org/10.1080/00031305.1995.10476109
3. J. Avérous, M. Meste, Skewness for multivariate distributions: two approaches, *Ann. Stat.*, **25** (1997), 1984–1997. https://doi.org/10.1214/aos/1069362381
4. M. Ekström, S. R. Jammalamadaka, A general measure of skewness, *Stat. Probabil. Lett.*, **82** (2012), 1559–1568. https://doi.org/10.1016/j.spl.2012.04.011
5. K. V. Mardia, Measures of multivariate skewness and kurtosis with applications, *Biometrika*, **57** (1970), 519–530. https://doi.org/10.1093/biomet/57.3.519
6. S. N. Roy, On a heuristic method of test construction and its use in multivariate analysis, *Ann. Math. Statist*, **24** (1953), 220–238. https://doi.org/10.1214/aoms/1177729029
7. J. F.Malkovich, A. A. Afifi, On tests for multivariate normality, *J. Am. Stat. Assoc.*, **68** (1973), 176–179. https://doi.org/10.1080/01621459.1973.10481358
8. K. S. Song, Rényi information, loglikelihood and an intrinsic distribution measure, *J. Stat. Plann. Infer.*, **93** (2001), 51–69. https://doi.org/10.1016/S0378-3758(00)00169-5
9. N. Henze, On Mardia's kurtosis test for multivariate normality, *Commun. Stat.-Theor. M.*, **23** (1994), 1031–1045. https://doi.org/10.1080/03610929408831303
10. A. Azzalini, A. D. Valle, The multivariate skew-normal distribution, *Biometrika*, **83** (1996), 715–726. https://doi.org/10.1093/biomet/83.4.715
11. A. Rényi, On measures of entropy and information, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1** (1961), 547–561.
12. M. G. Genton, X. Liu, Moments of skew-normal random vectors and their quadratic forms, *Stat. Probabil. Lett.*, **51** (2003), 319–325. https://doi.org/10.1016/S0167-7152(00)00164-4
13. T. Kollo, Multivariate skewness and kurtosis measures with an application in ICA, *J. Multivariate Anal.*, **99** (2008), 2328–2338. https://doi.org/10.1016/j.jmva.2008.02.033
14. N. Balakrishnan, B. Scarpa, Multivariate measures of skewness for the skew-normal distribution, *J. Multivariate Anal.*, **104** (2012), 73–87. https://doi.org/10.1016/j.jmva.2011.06.017
15. W. Tian, C. Wang, T.wang, The multivariate extended skew normal distribution and its quadratic forms, in *Huynh, VN., Kreinovich, V., Sriboonchitta, S. (eds) Causal Inference in Econometrics*, Springer, Cham, 2016.

16. H. M. Kim, C. Kim, Moments of scale mixtures of skew-normal distributions and their quadratic forms, *Commun. Stat.-Theor. M.*, **46** (2017), 1117–1126. https://doi.org/10.1080/03610926.2015.1011339

17. S. R. Jammalamadaka, E. Taufer, G. H. Terdik, On multivariate skewness and kurtosis, *Sankhya A*, **83** (2021), 607–644. https://doi.org/10.1007/s13171-020-00211-6

18. M. Abdi, M. Madadi, N. Balakrishnan, A. Jamalizadeh, Family of mean-mixtures of multivariate normal distributions: Properties, inference and assessment of multivariate skewness, *J. Multivariate Anal.*, **181** (2021), 104679. https://doi.org/10.1016/j.jmva.2020.104679

19. R. B. Arellano-Valle, A. Azzalini, A formulation for continuous mixtures of multivariate normal distributions, *J. Multivariate Anal.*, **185** (2021), 104780. https://doi.org/10.1016/j.jmva.2021.104780

20. M. Amiri, N. Balakrishnan, Hessian and increasing-Hessian orderings of scale-shape mixtures of multivariate skew-normal distributions and applications, *J. Comput. Appl. Math.*, **402** (2022), 113801. https://doi.org/10.1016/j.cam.2021.113801

21. B. Zuo, C. Yin, Tail conditional expectations for generalized skew-elliptical distributions, *Probability in the Engineering and Informational Sciences*, **36** (2022), 500–513. https://doi.org/10.1017/S0269964820000674

22. M. E. Johnson, *Multivariate Statistical Simulation*, John Wiley & Sons, New York, 1987. https://doi.org/10.1002/9781118150740

23. K. T. Fang, Y. T. Zhang, *Generalized Multivariate Analysis*, Science Press & Springer, Beijing & Berlin, 1990.

24. A. Azzalini, A class of distributions which includes the normal ones, *Scand. J. Stat.*, **12** (1985), 171–178.

25. M. D. Branco, D. K. Dey, A general class of multivariate skew-elliptical distributions, *J. Multivariate Anal.*, **79** (2001), 99–113. https://doi.org/10.1006/jmva.2000.1960

26. A. Azzalini, An overview on the progeny of the skew-normal family - A personal perspective, *J. Multivariate Anal.*, **188** (2022), 104851. https://doi.org/10.1016/j.jmva.2021.104851

27. S. X. Lee, G. J. McLachlan, An overview of skew distributions in model-based clustering, *J. Multivariate Anal.*, **188** (2022), 104853. https://doi.org/10.1016/j.jmva.2021.104853

28. A. Azzalini, A. Capitanio, *The Skew-Normal and Related Families*, IMS monographs, Cambridge University Press, Cambridge, 2014. https://doi.org/10.1017/CBO9781139248891

29. E. Gómez, M. A. Gómez-Villegas, J. M. Marín, A survey on continuous elliptical vector distributions, *Revista Matemática Complutense*, **16** (2003), 345–361. https://doi.org/10.5209/rev_REMA.2003.v16.n1.16889

30. G. Terdik, *Multivariate Statistical Methods: Going Beyond the Linear*, Springer Nature, 2021. https://doi.org/10.1007/978-3-030-81392-5

31. M. G. Genton, *Skew-elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman & Hall/CRC Press, Boca Raton, 2004.

32. J. R. Schott, *Matrix Analysis for Statistics*, John Wiley & Sons, New York, 1997.

33. B. Zuo, C. Yin, Multivariate tail covariance for generalized skew-elliptical distributions, *J. Comput. Appl. Math.*, **410** (2022), 114210. https://doi.org/10.1016/j.cam.2022.114210

Appendix

Proof of Example 3.1. From Lemma 3.2,

$$h_R(r) = \frac{2\pi^{\frac{p+1}{2}}}{\Gamma\left(\frac{p+1}{2}\right)} r^p g^{(p+1)}\left(r^2\right) = \frac{2\Gamma\left(\frac{\nu+p+1}{2}\right)}{\nu^{\frac{p+1}{2}}\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} r^p \left(1+\frac{r^2}{\nu}\right)^{-\frac{\nu+p+1}{2}},$$

let $t = \frac{r^2}{\nu}$, $J = \frac{1}{2}\sqrt{\frac{\nu}{t}}$, then we obtain

$$\begin{aligned}
E(R) = \int_0^\infty r h_R(r)\,\mathrm{d}r &= \frac{2\Gamma\left(\frac{\nu+p+1}{2}\right)}{\nu^{\frac{p+1}{2}}\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty r^{p+1}\left(1+\frac{r^2}{\nu}\right)^{-\frac{\nu+p+1}{2}} \mathrm{d}r \\
&= \frac{\nu^{\frac{1}{2}}\Gamma\left(\frac{\nu+p+1}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty t^{\frac{p}{2}}(1+t)^{-\frac{\nu+p+1}{2}}\,\mathrm{d}t \\
&= \frac{\nu^{\frac{1}{2}}\Gamma\left(\frac{\nu+p+1}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} B\left(\frac{p}{2}+1, \frac{\nu-1}{2}\right) = \frac{\sqrt{\nu}\Gamma\left(\frac{\nu-1}{2}\right)\Gamma\left(\frac{p}{2}+1\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}, \ \nu > 1.
\end{aligned}$$

$$\begin{aligned}
E\left(R^2\right) = \int_0^\infty r^2 h_R(r)\,\mathrm{d}r &= \frac{2\Gamma\left(\frac{\nu+p+1}{2}\right)}{\nu^{\frac{p+1}{2}}\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty r^{p+2}\left(1+\frac{r^2}{\nu}\right)^{-\frac{\nu+p+1}{2}} \mathrm{d}r \\
&= \frac{\nu\Gamma\left(\frac{\nu+p+1}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty t^{\frac{p+1}{2}}(1+t)^{-\frac{\nu+p+1}{2}}\,\mathrm{d}t \\
&= \frac{\nu\Gamma\left(\frac{\nu+p+1}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} B\left(\frac{p+1}{2}+1, \frac{\nu}{2}-1\right) = \frac{\nu(p+1)}{\nu-2}, \ \nu > 2.
\end{aligned}$$

Similarly, the following analogy can be obtained:

$$E\left(R^3\right) = \frac{\nu^{\frac{3}{2}}\Gamma\left(\frac{\nu+p+1}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} B\left(\frac{p}{2}+2, \frac{\nu-1}{2}-1\right) = \frac{\nu^{\frac{3}{2}}\Gamma\left(\frac{\nu-1}{2}-1\right)\Gamma\left(\frac{p}{2}+2\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}, \ \nu > 3,$$

and

$$E\left(R^4\right) = \frac{\nu^2\Gamma\left(\frac{\nu+p+1}{2}\right)}{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} B\left(\frac{p+1}{2}+2, \frac{\nu}{2}-2\right) = \frac{\nu^2(p+3)(p+1)}{(\nu-2)(\nu-4)}, \ \nu > 4.$$

So we obtain

$$\frac{E(R)}{E(R_0)} = \sqrt{\frac{\nu}{2}}\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}, \ \nu > 1; \quad \frac{E\left(R^2\right)}{E\left(R_0^2\right)} = \frac{\nu}{\nu-2}, \ \nu > 2$$

and

$$\frac{E\left(R^{3}\right)}{E\left(R_{0}^{3}\right)}=\left(\frac{\nu}{2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu-1}{2}-1\right)}{\Gamma\left(\frac{\nu}{2}\right)},\ \nu>3;\quad \frac{E\left(R^{4}\right)}{E\left(R_{0}^{4}\right)}=\frac{\nu^{2}}{(\nu-2)(\nu-4)},\ \nu>4.$$

□