应用概率统计　第二十四卷
第六期　2008年12月

Chinese Journal of Applied Probability
and Statistics Vol.24 No.6 Dec. 2008

Moments of the Skew Normal Random Vectors in the General Case

Fang Biqi

(*Institute of Applied Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, 100190*)

Abstract

In this paper, we derived the moments of the random vectors with the skew normal distributions and their quadratic forms in the general case. As an application, the measures of multivariate skewness and kurtosis are calculated.

Keywords: Skew normal distribution, quadratic form, skewness, kurtosis.

AMS Subject Classification: 62H99.

§1. Introduction

The skew normal distributions are the extensions of the normal distributions and are useful in modelling the data presenting skewness. A p-dimensional random vector $\mathbf{z}$ is called to have a skew normal distribution, if it has probability density function

$$\Phi(\lambda+\alpha'(\mathbf{z}-\mu))\phi_p((\mathbf{z}-\mu)'\Omega^{-1}(\mathbf{z}-\mu))|\Omega|^{-1/2}/\Phi(\lambda/c_0), \tag{1.1}$$

where $\lambda\in R^1$, $\alpha,\mu\in R^p$, $\Omega>0$ is $p\times p$, $c_0=(1+\alpha'\Omega\alpha)^{1/2}$, $\phi_p(x)=(2\pi)^{-p/2}\exp(-x/2)$, $\Phi(\cdot)$ is the distribution function of standard normal $N(0,1)$, see Arnold and Beaver (2000, Eq. (4.9)). It is the skew elliptical distribution with normal density generator and denoted by $S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$, see Fang (2003) and the multivariate closed skew normal distribution, $\mathrm{CSN}_{p,q}(\xi,\Omega,\alpha',-\lambda,1)$ with $q=1$ in González-Farías, Domínguez-Molina and Gupta (2004). A special case of (1.1) that $\lambda=0$ was studied by Azzalini and Dalla valle (1996), Azzalini and Capitanio (1999) and Genton, He and Liu (2001), among others. The case of $\lambda=0$ in (1.1) was generalized to the skew elliptical distribution in Branco and Dey (2001), the generalized skew-elliptical distribution (GSE) in Genton and Loperfido (2005). In this paper we confine to the investigation of the skew normal distributions in the form of (1.1).

In many applications, the restriction $\lambda=0$ is not suitable and the more general form with unrestricted λ must be considered. For example, Arnold and Beaver (2000) showed

Received August 5, 2005. Revised October 6, 2007.

by the likelihood ratio test in a real example that the model with all parameters in (1.1) provides the best fit among the three models: the normal distribution $N(\mu, \Omega)$, the skew normal distribution with $\lambda = 0$ and the skew normal distribution with unrestricted λ. The skew normal distribution can arise in the situations in which the observed normal random variables are truncated by some hidden variable. Specifically, let

$$\delta = \Omega\alpha/c_0, \qquad \Omega^* = \begin{pmatrix} 1 & \delta' \\ \delta & \Omega \end{pmatrix}. \tag{1.2}$$

Suppose $(x_0, \mathbf{x}') \sim N_{p+1}(0, \Omega^*)$ and

$$\mathbf{z} \stackrel{\mathrm{d}}{=} \mu + (\mathbf{x} | x_0 + \lambda/c_0 > 0). \tag{1.3}$$

Then $\mathbf{z} \sim S_p(\mu, \Omega, \lambda, \alpha; \phi_{p+1})$, see Arnold and Beaver (2000, p. 27) and Fang (2003). The parameter λ appears in the case that the truncation on the hidden variable x_0 is at a level not equal to its mean. Other situations in which the skew normal distribution with non-zero λ arises include the selected model in Copas and Li (1997). Let y be the response variable of interest and $\mathbf{x}$ the vector of covariates. The main model is

$$y = \beta'\mathbf{x} + \sigma\epsilon_1, \tag{1.4}$$

supplemented by a 'selection equation'

$$z = \gamma'\mathbf{x} + \epsilon_2. \tag{1.5}$$

The joint distribution of ϵ_1 and ϵ_2 is standard bivariate normal with correlation coefficient ρ. Then by (1.3) the resulting distribution of y given $\mathbf{x}$ and $z > 0$ is the skew normal $S_1(\beta'\mathbf{x}, \sigma^2, (1-\rho^2)^{-1/2}\gamma'\mathbf{x}, \rho(1-\rho^2)^{-1/2}\sigma^{-1}; \phi_2)$ in the notation here. Applications of this model were made to missing data and comparative trials.

The general form of (1.1) is important in the theoretical study of the skew normal distributions. The family of the skew normal distributions is closed under linear transformation, marginalization and conditioning. For the closure under conditioning, the general form with parameter λ is essential, see Arnold and Beaver (2000, Eq. (4.8)). More closure properties are achieved by the multivariate closed skew normal distribution for the sum and the joint distribution of independent random vectors in González-Farías, Domínguez-Molina and Gupta (2004). Some properties, for example, the invariance of the distribution of the centralized quadratic forms, obtained for the case $\lambda = 0$ do not hold if $\lambda \neq 0$. Some statistics are formed by centralized quadratic forms, which can be regarded as the quadratic forms of random vectors with location zero. If $\lambda = 0$, then the distribution of the centralized quadratic form of the skew normal distribution does not depend on the skewness parameter α and is the same as that under normality, see Azzalini and Capitanio (1999, Proposition 7). This property has been generalized in Genton and Loperfido (2005) for the generalized skew-elliptical distributions that the distribution of an even function of GSE random vector with location zero does not depend on the specific skewing function, which in the case of (1.1) is $\Phi(\alpha'\mathbf{z})$. However, in the case that $\lambda \neq 0$, the distribution of the centralized quadratic form will depend on both skewness parameters λ and α. Moreover, many statistics involve random vectors with non-zero location for their non-null distributions and the invariance argument can not be applied to obtain their distributions even if $\lambda = 0$. Examples of the statistics investigated related to these issues for the skew elliptical distributions include the F statistic (Fang, 2005), the Hotelling's T^2 statistic (Fang, 2006) and the four statistics for testing the linear restriction on the location parameters in the family of the matrix variate skew elliptical distributions, the likelihood ratio, the generalized T_0^2 statistic in Hotelling (1947) and Lawley (1938), statistic V in Pillai (1955), and the largest latent root in Roy (1953), see Muirhead (1982, p. 441), Fang (2005).

In this paper we extend the investigation of Genton, He and Li (2001) for the moments of the skew normal distribution with probability density function (1.1) restricted to $\lambda = 0$ into the general case. In Section 2, we obtain the first four moments of $\mathbf{z} \sim S_p(\mu, \Omega, \lambda, \alpha; \phi_{p+1})$ and the first two moments of its quadratic forms. In Section 3, the expressions of the measures of the multivariate skewness and kurtosis as defined by Mardia (1970) for the skew normal population are obtained. Estimates of the population skewness and kurtosis are given for a real data set.

§2. Moments

We shall omit the proofs of the results in this section, which are based on the properties of the matrix derivatives (Graham, 1981) and the communication matrix (Magnus and Neudecker, 1979).

The moment generating function of the distribution (1.1) with $\mu = 0$ is

$$M(t) = \Phi(\lambda/c_0 + \mathbf{t}'\delta)\exp[(\mathbf{t}'\Omega\mathbf{t})/2]/\Phi(\lambda/c_0), \tag{2.1}$$

where $c_0 = (1+\alpha'\Omega\alpha)^{1/2}$, $\delta = \Omega\alpha/c_0$, see Arnold and Beaver (2000, Eq. (4.5)) for the case $\Omega = I$. Let $J_1 = \lambda/c_0 + \mathbf{t}'\delta$, $J_2 = (\mathbf{t}'\Omega\mathbf{t})/2$, $\zeta_1 = \phi(J_1)/\Phi(J_1)$, where $\phi(\cdot)$ is the probability density function of the standard univariate normal distribution. To obtain the moments of the skew normal distribution, we calculate the partial derivatives of $M(\mathbf{t})$ as follows.

$$\frac{\partial M(\mathbf{t})}{\partial \mathbf{t}} = \exp(J_2)[\Omega\mathbf{t}\Phi(J_1) + \delta\phi(J_1)]/\Phi(\lambda/c_0), \tag{2.2}$$

$$\frac{\partial^2 M(\mathbf{t})}{\partial \mathbf{t}\partial \mathbf{t}'} = \exp(J_2)\{\Phi(J_1)[\Omega + \Omega\mathbf{t}\mathbf{t}'\Omega] + \phi(J_1)[\Omega\mathbf{t}\delta' + \delta\mathbf{t}'\Omega - J_1\delta\delta']\}/\Phi(\lambda/c_0), \tag{2.3}$$

$$\begin{aligned}\frac{\partial^3 M(\mathbf{t})}{\partial\mathbf{t}\partial\mathbf{t}'\partial\mathbf{t}} &= \exp(J_2)\{\Phi(J_1)[\mathrm{vec}(\Omega)\mathbf{t}'\Omega+\Omega\otimes(\Omega\mathbf{t})+\Omega\mathbf{t}\otimes\Omega+(\Omega\mathbf{t})\otimes(\Omega\mathbf{t}\mathbf{t}'\Omega)]\\ &\quad+\phi(J_1)[\delta\otimes\Omega+\delta\otimes(\Omega\mathbf{t}\mathbf{t}'\Omega)+\mathrm{vec}(\Omega)\delta'+\Omega\otimes\delta\\ &\quad+(\Omega\mathbf{t})\otimes(\Omega\mathbf{t}\delta')+(\Omega\mathbf{t})\otimes(\delta\mathbf{t}'\Omega)-J_1\delta\otimes(\Omega\mathbf{t}\delta')-J_1\delta\otimes(\delta\mathbf{t}'\Omega)\\ &\quad-J_1(\Omega\mathbf{t}\otimes\delta\delta')+(J_1^2-1)\delta\otimes\delta'\otimes\delta]\}/\Phi(\lambda/c_0),\end{aligned}\tag{2.4}$$

$$\begin{aligned}\frac{\partial^4 M(\mathbf{t})}{\partial\mathbf{t}\partial\mathbf{t}'\partial\mathbf{t}\partial\mathbf{t}'}|_{\mathbf{t}=0} &= \exp(J_2)\{\Phi(J_1)[(I+K_{pp})(\Omega\otimes\Omega)+\mathrm{vec}(\Omega)\mathrm{vec}(\Omega)']\\ &\quad+J_1\phi(J_1)[-(I+K_{pp})(\Omega\otimes\delta\otimes\delta')-(I+K_{pp})(\delta\otimes\delta'\otimes\Omega)\\ &\quad-(\delta\otimes\delta)\mathrm{vec}(\Omega)'-\mathrm{vec}(\Omega)(\delta'\otimes\delta')\\ &\quad+(3-J_1^2)\delta\otimes\delta'\otimes\delta\otimes\delta']\}/\Phi(\lambda/c_0)|_{\mathbf{t}=0},\end{aligned}\tag{2.5}$$

where in (2.5) the terms which are equal to zeros at $\mathbf{t}=0$ are omitted.

Denote the moments of $\mathbf{z}\sim S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$ by $M_1=\mathsf{E}(\mathbf{z})$, $M_2=\mathsf{E}(\mathbf{z}\mathbf{z}')$, $M_3=\mathsf{E}(\mathbf{z}\otimes\mathbf{z}'\otimes\mathbf{z})$, $M_4=\mathsf{E}(\mathbf{z}\otimes\mathbf{z}'\otimes\mathbf{z}\otimes\mathbf{z}')$. Letting $\mathbf{t}=0$ in (2.2), (2.3), (2.4) and (2.5), we obtain the first four moments of the skew normal distribution (with location $\mu=0$) stated in the following lemma. For notation simplicity we use J_1 to denote its value at $\mathbf{t}=0$.

Lemma 2.1 Assume $\mathbf{z}\sim S_p(0,\Omega,\lambda,\alpha;\phi_{p+1})$. Then the first four moments of $\mathbf{z}$ are

$$\begin{aligned}M_1 &= \zeta_1\delta,\\ M_2 &= \Omega-J_1\zeta_1\delta\delta',\\ M_3 &= \zeta_1[\delta\otimes\Omega+\mathrm{vec}(\Omega)\delta'+\Omega\otimes\delta+(J_1^2-1)\delta\otimes\delta'\otimes\delta],\\ M_4 &= (I+K_{pp})(\Omega\otimes\Omega)+\mathrm{vec}(\Omega)\mathrm{vec}(\Omega)'\\ &\quad+J_1\zeta_1\{-(I+K_{pp})[\Omega\otimes(\delta\delta')]-(I+K_{pp})[(\delta\delta')\otimes\Omega]\\ &\quad-(\delta\otimes\delta)\mathrm{vec}(\Omega)'-\mathrm{vec}(\Omega)(\delta'\otimes\delta')\}+J_1(3-J_1^2)\zeta_1(\delta\otimes\delta'\otimes\delta\otimes\delta'),\end{aligned}$$

where $c_0=(1+\alpha'\Omega\alpha)^{1/2}$, $\delta=\Omega\alpha/c_0$, $J_1=\lambda/c_0$, $\zeta_1=\phi(J_1)/\Phi(J_1)$.

From Lemma 2.1 we obtain the moments in the case $\mu\neq0$ as stated in the following proposition.

Proposition 2.1 Assume $\mathbf{z}\sim S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$. Then the first four moments of $\mathbf{z}$ are

$$\begin{aligned}M_1 &= \zeta_1\delta+\mu,\\ M_2 &= \Omega-J_1\zeta_1\delta\delta'+\zeta_1(\mu\delta'+\delta\mu')+\mu\mu',\\ M_3 &= \zeta_1[\delta\otimes\Omega+\mathrm{vec}(\Omega)\delta'+\Omega\otimes\delta+(J_1^2-1)\delta\otimes\delta'\otimes\delta]\\ &\quad+\mu\otimes\Omega+\Omega\otimes\mu+\mathrm{vec}(\Omega)\otimes\mu'-J_1\zeta_1[\mu\otimes\delta'\otimes\delta+\delta\otimes\delta'\otimes\mu+\delta\otimes\mu'\otimes\delta]\\ &\quad+\zeta_1[\delta\otimes\mu'\otimes\mu+\mu\otimes\delta'\otimes\mu+\mu\otimes\mu'\otimes\delta]+\mu\otimes\mu'\otimes\mu,\end{aligned}$$

$$
\begin{aligned}
M_4 \ = \ & \Omega\otimes\Omega+\mathrm{vec}(\Omega)\mathrm{vec}(\Omega)'+K_{pp}(\Omega\otimes\Omega)\\
&+J_1\zeta_1\{-(I+K_{pp})[\Omega\otimes(\delta\delta')]-(I+K_{pp})[(\delta\delta')\otimes\Omega]\\
&-(\delta\otimes\delta)\mathrm{vec}(\Omega)'-\mathrm{vec}(\Omega)(\delta'\otimes\delta')\}+J_1(3-J_1^2)\zeta_1(\delta\otimes\delta'\otimes\delta\otimes\delta')\\
&+\zeta_1[\mu\otimes\delta'\otimes\Omega+\mu\otimes\delta\otimes\mathrm{vec}(\Omega)'+\mu\otimes\Omega\otimes\delta'+(J_1^2-1)\mu\otimes\delta'\otimes\delta\otimes\delta')\\
&+\delta'\otimes\Omega\otimes\mu+\delta\otimes\mathrm{vec}(\Omega)'\otimes\mu+\Omega\otimes\delta'\otimes\mu+(J_1^2-1)\delta'\otimes\delta\otimes\delta'\otimes\mu\\
&+\mu'\otimes\delta\otimes\Omega+\mu'\otimes\mathrm{vec}(\Omega)\otimes\delta'+\mu'\otimes\Omega\otimes\delta+(J_1^2-1)\mu'\otimes\delta\otimes\delta'\otimes\delta\\
&+\delta\otimes\Omega\otimes\mu'+\mathrm{vec}(\Omega)\otimes\delta'\otimes\mu'+\Omega\otimes\delta\otimes\mu'+(J_1^2-1)\delta\otimes\delta'\otimes\delta\otimes\mu']\\
&+\mu\otimes\mu'\otimes\Omega+\Omega\otimes\mu\otimes\mu'+\mu'\otimes\Omega\otimes\mu+\mu\otimes\Omega\otimes\mu'\\
&-J_1\zeta_1[\mu\otimes\mu'\otimes\delta\otimes\delta'+\delta\otimes\delta'\otimes\mu\otimes\mu'+\delta\otimes\mu'\otimes\mu\otimes\delta'+\mu\otimes\delta'\otimes\delta\otimes\mu']\\
&+\mathrm{vec}(\Omega)\otimes\mu'\otimes\mu'+\mathrm{vec}(\Omega)'\otimes\mu\otimes\mu-J_1\zeta_1[\delta\otimes\mu'\otimes\delta\otimes\mu'+\mu\otimes\delta'\otimes\mu\otimes\delta')]\\
&+\zeta_1[\delta\otimes\mu'\otimes\mu\otimes\mu'+\mu\otimes\mu'\otimes\mu\otimes\delta'+\mu\otimes\delta'\otimes\mu\otimes\mu'+\mu\otimes\mu'\otimes\delta\otimes\mu']\\
&+\mu\otimes\mu'\otimes\mu\otimes\mu'. \qquad (2.6)
\end{aligned}
$$

Applying the formulas $\mathsf{E}(\mathbf{z}'A\mathbf{z})=\mathrm{tr}[A\mathsf{E}(\mathbf{z}\mathbf{z}')]$ and $\mathsf{E}[(\mathbf{z}'A\mathbf{z})(\mathbf{z}'B\mathbf{z})]=\mathrm{tr}[(A\otimes B)\mathsf{E}(\mathbf{z}\otimes\mathbf{z}'\otimes\mathbf{z}\otimes\mathbf{z}')]$ and using formulas in Magnus and Neudecker (1979, Theorem 3.1), we obtain the first two moments of the quadratic form for $\mathbf{z}$ in Proposition 2.2.

Proposition 2.2 Assume $\mathbf{z}\sim S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$. Let A, B be two $p\times p$ symmetric matrices. Then

$$
\begin{aligned}
\mathsf{E}(\mathbf{z}'A\mathbf{z}) \ =\ & \mathrm{tr}(A\Omega)+\mu'A\mu+2\zeta_1\mu'A\delta-J_1\zeta_1\delta'A\delta, \qquad (2.7)\\
\mathsf{E}[(\mathbf{z}'A\mathbf{z})(\mathbf{z}'B\mathbf{z})] \ =\ & 2\mathrm{tr}(A\Omega B\Omega)+\mathrm{tr}(A\Omega)\mathrm{tr}(B\Omega)\\
&+J_1\zeta_1[-\mathrm{tr}(A\Omega)\delta'B\delta-\mathrm{tr}(B\Omega)\delta'A\delta-4\delta'A\Omega B\delta+(3-J_1^2)\delta'A\delta\delta'B\delta]\\
&+2\zeta_1[2\mu'B\Omega A\delta+2\mu'A\Omega B\delta+\mathrm{tr}(A\Omega)\mu'B\delta+\mathrm{tr}(B\Omega)\mu'A\delta]\\
&+2\zeta_1(J_1^2-1)[\mu'B\delta\delta'A\delta+\mu'A\delta\delta'B\delta]\\
&+\mathrm{tr}(B\Omega)\mu'A\mu+\mathrm{tr}(A\Omega)\mu'B\mu+4\mu'A\Omega B\Omega\mu\\
&-J_1\zeta_1[\mu'A\mu\delta'B\delta+4\mu'A\delta\mu'B\delta+\mu'B\mu\delta'A\delta]\\
&+2\zeta_1[\mu'A\delta\mu'B\mu+\mu'A\mu\mu'B\delta]+\mu'A\mu\mu'B\mu, \qquad (2.8)\\
\mathsf{Cov}\,(\mathbf{z}'A\mathbf{z},\mathbf{z}'B\mathbf{z}) \ =\ & 2\mathrm{tr}(A\Omega B\Omega)+2\mu'(A\Omega B+B\Omega A)(\mu+2\zeta_1\delta)\\
&+2\zeta_1(J_1\zeta_1+J_1^2-1)(\delta'A\delta\mu'B\delta+\mu'A\delta\delta'B\delta)\\
&-4(\zeta_1^2+J_1\zeta_1)\mu'A\delta\mu'B\delta\\
&-4J_1\zeta_1\delta'A\Omega B\delta+J_1\zeta_1(-J_1\zeta_1+3-J_1^2)\delta'A\delta\delta'B\delta. \qquad (2.9)
\end{aligned}
$$

With $A=B$ in Proposition 2.2, we obtain $\mathsf{Var}\,(\mathbf{z}'A\mathbf{z})$. The mean and covariance of one quadratic form of the skew normal random vector $\mathbf{z}\sim S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$ were obtained in Fang (2005) from the moment generating function of the quadratic form. The first two moments of the skew normal distribution with location zero were obtained in Arnold and Beaver (2000). If $\lambda=0$ so that $J_1=0$ and $\zeta_1=(2/\pi)^{1/2}$, then the formulas in Proposition 2.1 and Proposition 2.2 reduce to those in Genton, He and Liu (2001).

§3. Application

Measures of multivariate skewness β_{1p} and kurtosis β_{2p} for a p-dimensional random vector $\mathbf{x}$ with mean μ and covariance Σ were defined in Mardia (1970, Eqs. (2.19), (3.5)) as

$$\beta_{1p}=\sum_{r,s,t}\sum_{r',s',t'}\sigma^{rr'}\sigma^{ss'}\sigma^{tt'}\mu_{111}^{(rst)}\mu_{111}^{(r's't')}, \tag{3.1}$$

$$\beta_{2p}=\mathsf{E}\{[(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)]^2\}, \tag{3.2}$$

where $\mu_{111}^{(rst)}=\mathsf{E}[(x_r-\mu_r)(x_s-\mu_s)(x_t-\mu_t)]$, $\sigma^{rr'}$ is the (r,r')-th element of Σ^{-1}.

Proposition 3.1 The measures of multivariate skewness and kurtosis of the distribution $S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$ are β_{1p} given by (3.1) and

$$\begin{aligned}\beta_{2p} \;=\; & 2\mathrm{tr}(\Sigma^{-1}\Omega\Sigma^{-1}\Omega)+[\mathrm{tr}(\Sigma^{-1}\Omega)]^2-2\zeta_1(J_1+\zeta_1)\mathrm{tr}(\Sigma^{-1}\Omega)\delta'\Sigma^{-1}\delta\\ &+\zeta_1(-3\zeta_1^3-6J_1\zeta_1^2-4J_1^2\zeta_1+4\zeta_1+3J_1-J_1^3)(\delta'\Sigma^{-1}\delta)^2\\ &-4\zeta_1(J_1+\zeta_1)\delta'\Sigma^{-1}\Omega\Sigma^{-1}\delta,\end{aligned} \tag{3.3}$$

where

$$\mu_{111}^{(rst)}=\delta_r\delta_t\delta_s\zeta_1(3J_1\zeta_1+2\zeta_1^2+J_1^2-1), \tag{3.4}$$

$$\Sigma=\Omega-\zeta_1(J_1+\zeta_1)\delta\delta'. \tag{3.5}$$

Proof We obtain $\mathsf{Var}\,(\mathbf{z})=\Sigma$ given by (3.5) by Lemma 2.1. If $\mathbf{z}\sim S_p(\mu,\Omega,\lambda,\alpha;\phi_{p+1})$, then $\mathbf{z}-\mathsf{E}(\mathbf{z})\sim S_p(-\zeta_1\delta,\Omega,\lambda,\alpha;\phi_{p+1})$ by Proposition 2.1. Substituting μ by $-\zeta_1\delta$ in M_3 in Proposition 2.1, we obtain $\mu_{111}^{(rst)}$ given by (3.4). Letting $B=A$, we obtain from Proposition 2.2,

$$\begin{aligned}\mathsf{E}[(\mathbf{z}'A\mathbf{z})^2] \;=\; & 2\mathrm{tr}(A\Omega A\Omega)+[\mathrm{tr}(A\Omega)]^2\\ &+J_1\zeta_1[-2\mathrm{tr}(A\Omega)\delta'A\delta-4\delta'A\Omega A\delta+(3-J_1^2)(\delta'A\delta)^2]\\ &+\zeta_1[8\mu'A\Omega A\delta+4\mathrm{tr}(A\Omega)\mu'A\delta+4(J_1^2-1)\mu'A\delta\delta'A\delta]\\ &+2\mathrm{tr}(A\Omega)\mu'A\mu+4\mu'A\Omega A\mu-J_1\zeta_1[2\mu'A\mu\delta'A\delta+4(\mu'A\delta)^2]\\ &+4\zeta_1\mu'A\delta\mu'A\mu+(\mu'A\mu)^2.\end{aligned} \tag{3.6}$$

Substituting μ by $-\zeta_1\delta$ and A by Σ^{-1} in (3.6)), we obtain (3.3). □

As an example, we use the estimates of the parameters obtained in Arnold and Beaver (2000) to estimate the multivariate skewness and kurtosis of a skew normal population. The data set reported by Cook and Weisberg (1994) for 202 athletes at the Australian Institute of Sport contains their heights and weights. Arnold and Beaver (2000) assumed the person's height and weight have joint distribution $S_2(\mu,\Sigma,\lambda_0,\Sigma^{-1/2}(\lambda_1,\lambda_2)';\phi_3)$ in the notation here. They obtained maximum likelihood estimates of the parameters for three models: the model under normality with $\lambda_0=0$, $\lambda_1=0$, $\lambda_2=0$; the model with $\lambda_0=0$; and the model with all parameters. Note Σ corresponds to Ω in Proposition 3.1. Substituting these estimates into β_{1p} and β_{2p} in Proposition 3.1, we obtain their estimates $\widehat{\beta}_{1p}$ and $\widehat{\beta}_{2p}$. In the first model the distribution of $\mathbf{z}$ is normal and $\beta_{1p}=0$, $\beta_{2p}=8$, which do not depend on the parameters. We obtain $\widehat{\beta}_{1p}=0.4917$, $\widehat{\beta}_{2p}=8.5449$ in the second model; $\widehat{\beta}_{1p}=0.7665$, $\widehat{\beta}_{2p}=9.1655$ in the third model. Using the data of the heights and weights of the 202 athletes, we obtain the sample analogue of β_{1p} as $b_{1p}=1.6878$ and the sample analogue of β_{2p} as $b_{2p}=10.8103$, see Mardia (1970, Eq. (2.23) and Eq. (3.12)) for the definitions of b_{1p} and b_{2p}. Estimates of the population skewness and kurtosis in the model with unrestricted λ_0 are the closest to the sample analogues among the three models. This is consistent with the conclusion in Arnold and Beaver (2000) that this model provides a best fit among the three models by the likelihood ratio test. Testing multivariate skewness and kurtosis can be made by using the statistics formed by the sample skewness and kurtosis in Mardia (1970). To test $\beta_{1p}=0$, a test statistic is $A=nb_{1p}/6$ and has the χ^2 distribution with $p(p+1)(p+2)/6$ degrees of freedom under hypothesis asymptotically, see Mardia (1970, Eq. (2.26)). To test $\beta_{2p}=p(p+2)$ which is the kurtosis under normality, a test statistic is $B^*=[b_{2p}-\{p(p+2)(n-1)/(n+1)\}]/\{8p(p+2)/n\}^{1/2}$ and has standard normal distribution under hypothesis asymptotically, see Mardia (1970, Eq. (3.20)). Using the data in this example, we obtain $A=56.8226$ and $B^*=5.1328$. Both of these values are highly significant. The rejection of the hypotheses indicates skewness and kurtosis in the distribution of the data.

§4. Discussion

The first four moments of the skew normal distributions in the general case that $\lambda\neq 0$ are given in Proposition 2.1, by which the first two moments of their quadratic forms are derived in Proposition 2.2. These formulas generalize the results in the literature. See remark below Proposition 2.2. Though these formulas involve many terms, their numerical calculation are straightforward and can be implemented by computer instantly. The distributions of the statistics of the skew normal vectors are usually very complex as the results of the presence of the skewness parameters and the formulas of moments can then be used to investigate their properties. Proposition 3.1 presents formulas for the calculation of the measures of multivariate skewness and kurtosis. They can be estimated when the estimates of the parameters are available and then used for the inference of the data as shown in the example in Section 3. However, the estimation of the skewness parameters for the skew normal distributions is usually not easy, see, for example, Azzalini and Capitanio (1999). The sample analogues of the multivariate skewness and kurtosis provide a simple way to investigate the shape of real data. Other statistical analyses are of course applicable depending on the purposes and methods available to the statistician. Further extension of the formulas of moments can be made for the more general distributions, for example, the skew elliptical distributions and the CSN distributions. The moments of the skew elliptical distributions can be calculated from those of the skew normal distribution by Fang (2003, Eq. (20)) if $\lambda = 0$ and from those with $\lambda = 0$ by the meathod in Fang (2003, Proposition 3) if $\lambda \neq 0$. The moments of the CSN distribution can be obtained without essential difficulties by the moment generating function in González-Farías, Domínguez-Molina and Gupta (2004), though the calculation is tedious.

References

[1] Arnold, B.C., Beaver, R.J., Hidden truncation models, *Sankhyā*, **62A**(2000), 23–35.

[2] Azzalini, A., Capitanio, A., Statistical applications of the multivariate skew normal distribution, *J. R. Statist. Soc. B*, **61**(1999), 579–602.

[3] Azzalini, A., Dalla Valle, A., The multivariate skew-normal distribution, *Biometrika*, **83**(1996), 715–726.

[4] Branco, M.D., Dey, D.K., A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, **79**(2001), 99–113.

[5] Cook, R.D., Weisberg, S., *An Introduction to Regression Graphics*, Wiley, New York, 1994.

[6] Copas, J.B., Li, H.G., Inference for non-random samples, *J. R. Statist. Soc. B*, **59**(1997), 55–99.

[7] Fang, B.Q., The skew elliptical distributions and their quadratic forms, *Journal of Multivariate Analysis*, **87**(2003), 298–314.

[8] Fang, B.Q., Noncentral quadratic forms of the skew elliptical variables, *Journal of Multivariate Analysis*, **95**(2005), 410–430.

[9] Fang, B.Q., The F statistic of the skew elliptical variables, *Chinese Journal of Applied Probability and Statistics*, **21**(2005), 197–206.

[10] Fang, B.Q., Invariant distributions of the multivariate tests in the skew elliptical model, *Statistical Methodology*, **2**(2005), 285–296.

[11] Fang, B.Q., Sample mean, covariance and T^2 statistic of the skew elliptical model, *Journal of Multivariate Analysis*, **97**(2006), 1675–1690.

[12] Genton, M.G., He, L., Liu, X., Moments of Skew-normal random vectors and their quadratic forms, *Statistics & Probability Letters*, **51**(2001), 319–325.

[13] Genton, M.G., Loperfido, N., Generalized skew-elliptical distributions and their quadratic forms, *Ann. Inst. Statist. Math.*, **57**(2005), 389–401.

[14] González-Farías, G., Domínguez-Molina, A., Gupta, A.K., Additive properties of skew normal random vectors, *Journal of Statistical Planning and Inference*, **126**(2004), 521–534.

[15] Graham, A., *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood Limited, Chichester, 1981.

[16] Hotelling, H., Multivariate quality control, illustrated by the air testing of sample bombsights, *Techniques of Statistical Analysis*, McGraw-Hill, New York, 1947, 111–184.

[17] Lawley, D.N., A generalization of Fisher's z test, *Biometrika*, **30**(1938), 180–187.

[18] Magnus, J.R., Neudecker, H., The commutation matrix: some properties and applications, *The Annals of Statistics*, **7**(1979), 381–394.

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

[20] Muirhead, R.J., *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

[21] Pillai, K.C.S., Some new test criteria in multivariate analysis, *Ann. Math. Statist.*, **26**(1955), 117–121.

[22] Roy, S.N., On a heuristic method of test construction and its use in multivariate analysis, *Ann. Math. Statist.*, **24**(1953), 220–228.

一般形式下斜正态随机向量的矩

方 碧 琪

(中国科学院数学与系统科学研究院应用数学研究所, 北京, 100190)

本文给出一般形式下斜正态随机向量及其平方型的矩公式. 作为应用, 计算出了斜正态随机向量的多元偏度和峰度.

关 键 词: 斜正态分布, 平方型, 偏度, 峰度.

学 科 分 类 号: O212.4.