JOURNAL OF MULTIVARIATE ANALYSIS **58**, 96–106 (1996)
ARTICLE NO. 0041

The Distribution of the Covariance Matrix for a Subset of Elliptical Distributions with Extension to Two Kurtosis Parameters

H. S. Steyn

University of Pretoria, Pretoria, South Africa

1. Introduction

To study the influence of deviations from normality the class of elliptical distributions, which introduces a single kurtosis parameter κ, is often

the moment generating function (m.g.f.) of the distribution of the sampling covariance matrix for a subclass of elliptical distributions is first derived and is then generalized to include distributions containing two kurtosis parameters.

A subclass of elliptical distributions, $E_m(\mathbf{0}, \mathbf{P}; \kappa)$, was introduced by the present author [4] by considering the following convergent series as cumulant generating function (c.g.f.),

$$K(\mathbf{t}) = \tfrac{1}{2}(\mathbf{t}'\mathbf{P}\mathbf{t}) + \tfrac{1}{2}\kappa[\tfrac{1}{2}(\mathbf{t}'\mathbf{P}\mathbf{t})]^2 + \sum_{r>2} A_r(\mathbf{t}'\mathbf{P}\mathbf{t})^r, \tag{1a}$$

with probability density function (p.d.f.),

$$e_m(\mathbf{x}; \mathbf{0}; \mathbf{P}, \kappa) = n_m(\mathbf{x}; \mathbf{0}, \mathbf{P})\{1 + \tfrac{1}{8}\kappa \mathbf{H}_2^m(Q)/1! + \text{terms of higher order in the polynomial } Q\}, \tag{1b}$$

where the matrix $\mathbf{P} = \mathbf{P}_{m \times m} = (\rho_{ij})$ is the population correlation matrix, $\mathbf{t}$ is an $m * 1$ generating vector, $A_3, A_4, ...,$ are constants, κ is a kurtosis parameter, while $n_m(\mathbf{x}; \mathbf{0}, \mathbf{P})$ is the standardized m-dimensional normal with quadratic form $Q = \mathbf{x}'\mathbf{P}^{-1}\mathbf{x}$. The definition of the polynomials $H_r^m(Q)$ are given in [4]. In particular, $\mathbf{H}_2^m(Q) = Q^2 - 2(m+2)\,Q + m(m+2)$. The subclass of elliptical distributions as defined by (1) contains quite a number

Received October 26, 1993; revised September 1995.
AMS 1990 subject classifications: 62H15, 62H10.
Key words and phrases: elliptical distributions, covariance matrix, moment generating functions, kurtosis parameter.

0047-259X/96 $18.00

of well-known members of the elliptical class, such as the multivariate normal, the contaminated multivariate normal and the multivariate t-distribution. An extension of $E_m(\mathbf{0}, \mathbf{P}; \kappa)$ to cases containing more than one kurtosis parameter is given in [4]. The m.g.f. as defined in (1a) will be used as starting point to derive the properties of the required sampling distributions of the covariance matrix for the case of one and two kurtosis parameters. An m.g.f. approach will be used in this paper. An alternative method for dealing with the case of one kurtosis parameter is given in a paper by Sutradhar and Ali [6].

2. The Distribution of the Covariance Matrix for One Kurtosis Parameter

To derive the m.g.f. of the covariance matrix for a random sample from the m-dimensional population with c.g.f. (1a), the procedure is as follows:

(i) The m.g.f. of the joint distribution of $x_i x_j$, $i \leqslant j$, $i = 1, 2, ..., m$; $j = 1, 2, ..., m$, is derived by first introducing a differential operator in terms of generating variables t_{ij} associated with $x_i x_j$ respectively and which operates on the relevant m.g.f. for the multinormal case.

(ii) The m.g.f. of the joint distribution of the second order moments about the origin is obtained from (i) by using straightforward probability theory and matrix differentiation.

(iii) The m.g.f. of the joint distribution of the elements of the covariance matrix is obtained by applying an orthogonal transformation to the distribution derived in (ii).

Starting with the case where $\mathbf{X}$ is multivariate normal with p.d.f. $n_m(\mathbf{x}; \mathbf{0}, \mathbf{P})$, the m.g.f. of the joint distribution of the squares and the products $x_i x_j$, $i \leqslant j$, $i, j = 1, 2, ..., m$, is given by

$$M_1(\mathbf{t}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{ \exp\left(\sum_{i \leqslant j = 1}^{m} t_{ij} x_i x_j \right) \right\} C \exp\{-\tfrac{1}{2} Q\} \prod_{i=1}^{m} dx_i, \quad (2)$$

where $\mathbf{t} = (t_{11}, t_{12}, ..., t_{mm})'$, $C = (2\pi)^{(1/2)m} |\mathbf{P}|^{-1/2}$ and $Q = \mathbf{x}'\mathbf{P}^{-1}\mathbf{x} = |\mathbf{P}|^{-1} \times \{\sum_{i=1}^{m} |\mathbf{P}_{ii}| \, x_i^{\,2} + 2 \sum_{i<j} |\mathbf{P}_{ij}| \, x_i x_j\}$. Differentiating (2) with respect to t_{ij} using

$$\mathscr{D}_{m:P} \text{ for the operator } |\mathbf{P}|^{-1} \left\{ \sum_{i=1}^{m} |\mathbf{P}_{ii}| \frac{\partial}{\partial t_{ii}} + 2 \sum_{i<j} |\mathbf{P}_{ij}| \frac{\partial}{\partial t_{ij}} \right\}, \quad (3)$$

it follows immediately that

$$\{\mathscr{D}_{m\,:\,\mathbf{P}}\}^s\, M_1(\mathbf{t}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{\exp\left(\sum_{i \leqslant j=1}^{m} t_{ij} x_i x_j\right)\right\} Q^s n_m(\mathbf{x};\, \mathbf{0},\, \mathbf{P}) \prod_{i=1}^{m} dx_i. \tag{4}$$

Applying the result in (4), the m.g.f. of the joint distribution of the squares x_i^2 and the products $x_i x_j$ where i and j take the values as indicated above, corresponding to the p.d.f. (1b) is given by

$$\begin{aligned} M_2(\mathbf{t}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{\exp\left(\sum_{i \leqslant j=1}^{m} t_{ij} x_i x_j\right)\right\} n_m(\mathbf{x};\, \mathbf{0},\, \mathbf{P}) \{1 + \tfrac{1}{8}\kappa \mathbf{H}_2^m(S)/1! \\ + \text{higher order terms in } Q\} \prod_{i=1}^{m} dx_i = \{1 + \tfrac{1}{8}\kappa \mathbf{H}_2^m(\mathscr{D}_{m\,:\,\mathbf{P}}) \\ + \text{terms of higher order in the operator } \mathscr{D}_{m\,:\,\mathbf{P}}\}\, M_1(\mathbf{t}). \end{aligned} \tag{5}$$

It is, however, well known (see [1, p. 160; 3, p. 89]) that

$$M_1(\mathbf{t}) = |\mathbf{P}^{-1}|^{1/2}\, |\mathbf{P}^{-1} - \mathbf{T}^*|^{-1/2},$$

where the matrix

$$\begin{aligned} \mathbf{T}^* = [(1+\delta_{ij})], \qquad t_{ij} = t_{ji}, \quad \delta_{ij} &= 1, \quad i = j \\ &= 0, \quad i \neq j. \end{aligned} \tag{6}$$

Thus, using (5) and (6), the m.g.f. of the joint distribution of $x_i x_j$, $i \leqslant j$, can be obtained by applying the operator $\mathscr{D}_{m\,:\,\mathbf{P}}$ on the m.g.f. $M_1(\mathbf{t})$. Considering n independent vectors $X = (X_{1s}, X_{2s}, ..., X_{ms})'$, $s = 1, 2, ..., n$ each with p.d.f. $e_m(\mathbf{x};\, \mathbf{0},\, \mathbf{P},\, \kappa)$ as in (1b), the m.g.f. of the sums of squares $\sum_{s=1}^{n} x_{is}^2$, $i = 1, 2, ..., m$, and the products $\sum_{s=1}^{n} x_{is} x_{js}$, $i \leqslant j$, is given by $[M_2(\mathbf{t})]^n$, so that the m.g.f. of the second-order moments about the mean, $\mu = 0$, is given by

$$M_3(\mathbf{t}) = \left[\left[M_2\left(\frac{\mathbf{t}}{n}\right)\right]\right]^n = \left[\left[\left\{1 + \frac{1}{8}\kappa \mathbf{H}_2^m(n\mathscr{D}_{m\,:\,\mathbf{P}}) + \cdots\right\} M_1\left(\frac{\mathbf{t}}{n}\right)\right]\right]^n, \tag{7}$$

where by (3)

$$\mathbf{H}_2^m(n\mathscr{D}_{m\,:\,\mathbf{P}}) = m(m+2) - 2(m+2)(n\mathscr{D}_{m\,:\,\mathbf{P}}) + (n\mathscr{D}_{m\,:\,\mathbf{P}})^2$$

and

$$M_1\left(\frac{\mathbf{t}}{n}\right) = |\mathbf{P}^{-1}|^{1/2} \left|\mathbf{P}^{-1} - \frac{1}{n}\mathbf{T}^*\right|^{-1/2}.$$

Writing now $\mathbf{Z}^* = \mathbf{P}^{-1} - (1/n)\,\mathbf{T}^*$, it follows on differentiation [1, p. 347] that

$$\frac{\partial}{\partial t_{ii}} |\mathbf{Z}^*|^{-1/2} = \frac{1}{n} |\mathbf{Z}^*|^{-1/2} |\mathbf{Z}^*_{ii}|\, |\mathbf{Z}^*|^{-1};$$
$$\frac{\partial}{\partial t_{ij}} |\mathbf{Z}^*|^{-1/2} = \frac{1}{n} |\mathbf{Z}^*|^{-1/2} |\mathbf{Z}^*_{ij}|\, |\mathbf{Z}^*|^{-1}, \tag{8}$$

where $|\mathbf{Z}^*_{ij}|$ is the cofactor of the element of the i th row and j-column in $|\mathbf{Z}^*|$. Also

$$\frac{\partial^2}{\partial t_{ij}\,\partial t_{kl}} |\mathbf{Z}^*|^{-1/2} = \frac{3}{n^2} |\mathbf{Z}^*|^{-1/2} |\mathbf{Z}^*_{ij}|\, |\mathbf{Z}^*_{kl}|\, |\mathbf{Z}^*|^{-2}$$
$$-\frac{1}{n^2} |\mathbf{Z}^*|^{-1/2} (|\mathbf{Z}^*_{kl,\,ij}| + |\mathbf{Z}^*_{kl,\,ji}|)\, |\mathbf{Z}^*|^{-1}, \tag{9}$$

where $|\mathbf{Z}^*_{kl,\,ij}|$ is the cofactor in $|\mathbf{Z}^*_{kl}|$ of the element in ith row and jth column of $|\mathbf{Z}^*|$.

It is clear from Eqs. (8) and (9) that the operator $\partial/\partial t_{ij}\,\partial t_{kl}$ leads to a factor $|\mathbf{Z}^*_{ij}|\,|\mathbf{Z}^*_{kl}|$ in the first term of the r.h.s. of (9) and to a factor $(|\mathbf{Z}^*_{kl,\,ij}| + |\mathbf{Z}^*_{kl,\,ji}|)$ in the second term. To write down the final form for $M_3(\mathbf{t})$ it is convenient at this stage to change the notation used in (2), where $i<j$. This notation gave rise to the factor 2 appearing in the second term of (3) and will have a further accumulating effect on the second-order operators. Using the subscripts p, q, r, s, all of which can take on the values $1, 2, 3, ..., m$, it now follows from (8) and (9), using (3), that

$$M_3(\mathbf{t}) = |\mathbf{P}^{-1}|^{(1/2)n} |\mathbf{Z}^*|^{-(1/2)n} \Bigg[1 + \tfrac{1}{8}\kappa m(m+2)$$
$$-\tfrac{1}{4}\kappa(m+2)|\mathbf{P}|^{-1} |\mathbf{Z}^*|^{-1} \sum_{p,\,q=1}^{m} |\mathbf{P}_{pq}|\, |\mathbf{Z}^*_{pq}|$$
$$+\tfrac{1}{8}\kappa\, |\mathbf{P}^{-2}|\, |\mathbf{Z}^*|^{-1} \sum_{p,\,q=1}^{m} \sum_{r,\,s=1}^{m} |\mathbf{P}_{pq}|\, |\mathbf{P}_{rs}|$$
$$\times \{3\, |\mathbf{Z}^*_{pq}|\, |\mathbf{Z}^*_{rs}|\, |\mathbf{Z}^*|^{-1} - (|\mathbf{Z}^*_{pq,\,rs}| + |\mathbf{Z}^*_{pq,\,sr}|)\} + \cdots \Bigg]^n, \tag{10}$$

where $|\mathbf{P}_{pq}| = |\mathbf{P}_{qp}|$, $|\mathbf{P}_{rs}| = |\mathbf{P}_{sr}|$, $|\mathbf{Z}^*_{pq}| = |\mathbf{Z}^*_{qp}|$, $|\mathbf{Z}^*_{rs}| = |\mathbf{Z}^*_{sr}|$, $\mathbf{Z}^* = \mathbf{P}^{-1} - (1/n)\,\mathbf{T}^*$.

The first factor in (10), that is, $|\mathbf{P}^{-1}|^{(1/2)n} |\mathbf{P}^{-1} - (1/n)\,\mathbf{T}^*|^{-(1/2)n} = |\mathbf{I}_m - (1/n)\,\mathbf{T}^*\mathbf{P}|^{-(1/2)n}$, is the well-known m.g.f. for the Wishart

$W_m(n, (1/n)\,\mathbf{P})$ distribution. It is clear that (10) reduces to this distribution when $\kappa = 0$.

To proceed from the m.g.f. of the second-order moments about the mean to the m.g.f. of the elements of the covariance matrix it is observed that the p.d.f. of $E_m(\mathbf{0}, \mathbf{P}; \kappa)$ is a function only of the quadratic form $Q = \mathbf{x}'\mathbf{P}^{-1}\mathbf{x}$. It follows, therefore, that the well-known procedure of orthogonal transformations (see, e.g., [3, p. 70]) is also applicable to this case. Thus, the m.g.f. $M_3()$ in (10) is the m.g.f. of the elements of the *covariance matrix*, $S = (s_{ij})$, with the parameters n now defined as $N-1$, where N is the sample size and where

$$s_{ij} = \frac{1}{n} \sum_{r=1}^{N} (x_{ir} - \bar{x}_i)(x_{jr} - \bar{x}_j).$$

The influence on (10), if terms of third and higher degrees in $(\mathbf{t}'\mathbf{P}\mathbf{t})$ were specified in the expressions (1a) and (1b), will be that (7) will include specific terms of higher order than the second in the differential operators. This will give rise to terms containing cofactors of the form $|\mathbf{Z}^*_{pq,\, rs,\, tu}|$ and lower order. However, the final form corresponding to (10) will contain the m.g.f. of the same Wishart distribution as first factor. It will be shown below that these terms have no influence on the standard errors of the elements of the covariance matrix.

The marginal distribution of any square submatrix $\mathbf{S}_l$ of $\mathbf{S}$ located on the diagonal of $\mathbf{S}$ follows from (10) but with m replaced by l the rank of $\mathbf{S}_l$. This is seen by substituting $t_{ij} = 0$ for $i, j > l$. Taking, for example, $l = 2$, the m.g.f. associated with the elements s_{11}, s_{12}, and s_{22} follows from (10) as

$$\begin{aligned} M_3(t_{11}, t_{12}, t_{22}) = \{A(t)\}^{-(1/2)\,n} \Bigg\{ & 1 + \kappa - 2\kappa \{A(t)\}^{-1} \\ & \times \left[1 - \frac{1}{n}(t_{11} + \rho_{12} t_{12} + t_{22}) \right] \\ & + \frac{3}{2} \kappa \{A(t)\}^{-2} \Bigg[1 - \frac{2}{n}(t_{11} + \rho_{12} t_{12} + t_{22}) \\ & + \frac{1}{n^2}(t_{11}^2 + \rho_{12}^2 t_{12}^2 + t_{22}^2 + 2t_{11}t_{22} + \rho_{12} t_{11} t_{12} + \rho_{12} t_{12} t_{22}) \Bigg] \\ & - \frac{1}{2} \kappa \{A(t)\}^{-1} + \cdots \Bigg\}^{n}, \end{aligned} \tag{11}$$

where $A(t) = 1 - (2/n)(t_{11} + \rho_{12} t_{12} + t_{22}) + (1/n^2)(1 - \rho_{12}^2)(4t_{11}t_{22} - t_{12}^2)$.

Hence, $E(s_{11})=1$ and $\text{var}(s_{11})=(1/n)(2+3\kappa)$, so that for an unstandardized population with variance σ_{11} $(=\sigma_1^2)$ follows that $\text{var}(s_{11})=(1/n)(2+3\kappa)\,\sigma_{11}^2$. $\text{Var}(s_{22})$ follows from $\text{var}(s_{11})$ by symmetry. The m.g.f. for the marginal distribution of s_{12} follows from (11) for $t_{11}=t_{22}=0$. Hence, $E(s_{12})=\rho_{12}$ and $\text{var}(s_{12})=[1+\rho_{12}^2+\kappa(1+2\rho_{12}^2]/n$, so that for the unstandardized case, $\text{var}(s_{12})=\sigma_{11}\sigma_{22}[1+\rho_{12}^2+\kappa(1+2\rho_{12}^2)]/n$. Similarly, putting $t_{22}=0$ in (11) to obtain the m.g.f. of the marginal distribution of (t_{11}, t_{12}) it follows from the coefficient of $t_{11}t_{12}$ that $\text{cov}(s_{11}, s_{12})=(1/n)(2\rho_{12}+3\kappa\rho_{12})\,\sigma_1^3\sigma_2^2$ for the unstandardized case. Further, for $t_{12}=0$ in (11), it follows for the unstandardized case that, $\text{cov}(s_{11}, s_{22})=(1/n)(2\rho_{12}^2+\kappa(1+2\rho_{12}^2)]\ \sigma_{11}\sigma_{22}$.

Rewriting the well-known expressions given in [3, pp. 41–42], for the variances and covariances of the *asymptotic distribution of the covariance matrix* under a multivariate elliptical model as

$$\text{cov}(s_{ij}, s_{kl})=\frac{1}{n}(\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk})+\frac{\kappa}{n}(\sigma_{ij}\sigma_{kl}+\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk}), \tag{12}$$

where, as usual, $\sigma_{ii}=\sigma_i^2$, $\sigma_{ij}=\rho_{ij}\sigma_i\sigma_j$, $i\neq j$, it is clear that the *exact* expressions for $\text{var}(s_{11})$, $\text{var}(s_{22})$, $\text{var}(s_{12})$, and $\text{cov}(s_{11}, s_{12})$ as derived from (11) for the subclass defined in (1), is the same as *asymptotic* expression (12). Clearly, it can be seen from the symmetrical property of (10) that (12) also gives the exact expressions for all elements of the covariance matrix.

3. The Distribution of the Covariance Matrix for Two Kurtosis Parameters

Dividing the random vector $\mathbf{X}'=(X_1, X_2, ..., X_m)$ into vectors $\mathbf{X}^{(1)}=(X_1, X_2, ..., X_h)'$ and $\mathbf{X}^{(2)}=(X_{h+1}, ..., X_m)'$ with a corresponding division of the generating variable $\mathbf{t}$ and of the matrix $\mathbf{P}$, and writing the m.g.f. of the multivariate normal in the form given in [4], Eq. (23), the c.g.f. for the case of two kurtosis parameters follows as

$$\begin{aligned} K(\mathbf{t}) &= \tfrac{1}{2}(\mathbf{t}^{(1)}+\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{t}^{(2)})'\,\mathbf{P}_{11}(\mathbf{t}^{(1)}+\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{t}^{(2)}) \\ &\quad+\tfrac{1}{2}\kappa_1[\tfrac{1}{2}(\mathbf{t}^{(1)}+\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{t}^{(2)})'\,\mathbf{P}_{11}(\mathbf{t}^{(1)}+\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{t}^{(2)})]^2 \\ &\quad+\tfrac{1}{2}(\mathbf{t}^{(2)\prime}\mathbf{P}_{22.1}\mathbf{t}^{(2)})+\tfrac{1}{2}\kappa_2[\tfrac{1}{2}(\mathbf{t}^{(2)\prime}\mathbf{P}_{22.1}\mathbf{t}^{(2)})]^2 \\ &\quad+\text{terms higher than fourth degree in } t_1, t_2, ..., t_m. \end{aligned} \tag{13}$$

where $\mathbf{P}_{22.1}=\mathbf{P}_{22}-\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}$. The associated quadratic forms are

$$Q_1=\mathbf{X}^{(1)\prime}\mathbf{P}_{11}^{-1}\mathbf{X}^{(1)}, \qquad Q_2=\mathbf{Y}^{(2)\prime}\mathbf{P}_{22.1}^{-1}\mathbf{Y}^{(2)}$$
$$\text{with}\quad \mathbf{Y}^{(2)}=\mathbf{X}^{(2)}-E(\mathbf{X}^{(2)}\mid \mathbf{X}^{(1)}=\mathbf{x}^{(1)}).$$

It is easily seen that the form of the c.g.f. (13) is such that m.g.f.'s of the distributions of submatrices or related conditional matrices of the sampling covariance matrix can be written down by using (10). For this purpose it is necessary to divide the sampling covariance matrix $\mathbf{S}$ into submatrices,

$$\mathbf{S}=\begin{bmatrix}\mathbf{S}_{11} & \mathbf{S}_{12}\\ \mathbf{S}_{12} & \mathbf{S}_{22}\end{bmatrix} \text{ and to define } \mathbf{S}_{22.1}=\mathbf{S}_{22}-\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}.$$

The relevant m.g.f.'s, as well as some properties, are

(a) Since the marginal distribution of $\mathbf{X}^{(1)}$ (obtained when $\mathbf{t}^{(2)}=0$ is $e_h(\mathbf{x}^{(1)}; \mathbf{0}\cdot\mathbf{P}_{11}, \kappa_1)$, the m.g.f. of the marginal distribution of the sampling covariance matrix S_{11} associated with the vector $\mathbf{X}^{(1)}$ is given by (10), where m is replaced by h, κ by κ_1, and $\mathbf{P}$ by $\mathbf{P}_{11}$. This means that the exact variances and covariances of the elements of $\mathbf{S}_{11}$ (the sampling covariance matrix corresponding to $\mathbf{P}_{11}$) can be written down by using (12).

(b) The distribution of the conditional sampling covariance $\mathbf{S}_{22.1}$ corresponding to the matrix $\mathbf{P}_{22.1}$ and associated with the conditional vector $\mathbf{Y}^{(2)}$ is distributed independently of $\mathbf{S}_{11}$. The m.g.f. of $\mathbf{S}_{22.1}$ follows directly from (10) by substituting $\mathbf{P}_{22.1}$ for $\mathbf{P}$, $m-h$ for m and κ_2 for κ and changing the relevant cofactors appearing in (10) accordingly.

(c) For the case of two kurtosis parameters an expression similar to m.g.f. (5) can be written down for the m.g.f. of the distribution of the elements of the matrix $[x_i x_j]$ of squares and products. This m.g.f. will be denoted by $M(\mathbf{t}; \kappa_1, \kappa_2)$ and is given by

$$\begin{aligned} M(\mathbf{t}; \kappa_1, \kappa_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} & \left\{\exp\left(\sum_{i\leqslant j=1}^{m} t_{ij}x_i x_j\right)\right\} n_h(\mathbf{x}^{(1)}; \mathbf{0}, \mathbf{P}_{11}) \\ & \times\{1+\tfrac{1}{8}\kappa_1 \mathbf{H}_2^s(Q_1)/1! + \text{terms of higher order}\} \\ & n_{m-h}(\mathbf{y}^{(2)}; \mathbf{0}, \mathbf{P}_{22.1})\{1+\tfrac{1}{8}\kappa_2\mathbf{H}_2^{m-h}(Q_2)/1! \\ & +\text{terms of higher order}\} \prod_1^m dx_i. \end{aligned} \tag{14}$$

To introduce the method consider the three-dimensional case when $m=3$ in (14) and where the random variables X_1, X_2, and X_3 are divided as (X_1, X_2) and X_3, so that $\mathbf{P}=(\rho_{ij})$, $i=1, 2, 3$; $j=1, 2, 3$;

$$\mathbf{P}_{11}=(\rho_{rs}),\ r=1, 2;\ s=1, 2;\ \rho_{pq}=1 \qquad \text{if} \quad p=q,$$

$$\mathbf{P}_{22.1}=1-\left[\rho_{12}\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}+\rho_{23}\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}\right]. \tag{15}$$

Further, substitute $t_{33}=0$, so that (14) reduces to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Bigg([\exp(t_{11}x_1^2+t_{12}x_1x_2+t_{22}x_2^2)]\, e_2(x_1,x_2;\mathbf{0},\mathbf{P}_{11},\kappa_1)$$
$$\times\Bigg\{\int_{-\infty}^{\infty}\exp\Bigg\{t_{13}\Bigg[x_1\Bigg(x_3+\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}x_1+\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}x_2\Bigg)$$
$$-\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}x_1^2-\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}x_1x_2\Bigg]\Bigg\}$$
$$\times\exp\Bigg\{t_{23}\Bigg[x_2\Bigg(x_3+\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}x_2+\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}x_1\Bigg)$$
$$-\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}x_1x_2-\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}x_2^2\Bigg]\Bigg\}$$
$$\times e_1(y_3;P_{22.1},\kappa_2)\,dx_3\Bigg\}\Bigg)\,dx_1\,dx_2, \tag{16}$$

where

$$Y_3=X_3-E(X_3\mid X_1=x_1,X_2=x_2)$$
$$=X_3+\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}x_2+\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}x_1.$$

Noting that (16) consists of a marginal factor and a conditional factor, it becomes clear that the function inside the large curly brackets is the m.g.f. of the joint conditional distribution of x_1x_3 and x_2x_3 given x_1^2, x_1x_2, and x_2^2. The following deductions are obvious:

(i) Putting $t_{23}=0$, it is seen that the conditional distribution of x_1x_3, given x_1^2, x_1x_2, and x_2^2, is univariate elliptical with mean and variance respectively equal to

$$\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}x_1x_2+\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}x_1^2,$$
$$\mathbf{P}_{22.1}x_1^2=\Bigg\{1-\Bigg[\rho_{13}\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}+\rho_{23}\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}\Bigg]\Bigg\}x_1^2. \tag{17}$$

Writing X_{3c} for the conditional X_3, similar expressions for the mean and the variance follow for the conditional distribution of x_2x_{3c} by interchanging the subscripts "1" and "2." It follows from (16) that the conditional product moments of x_1^2, x_1x_2, x_2^2 with x_1x_{3c} and with x_2x_{3c} are zero. The conditional product moment of x_1y_3 and x_2y_3 follows by integrating the coefficient of $t_{13}t_{23}$ in (16) and is given by

$$\mathbf{P}_{22.1}x_1x_2=\left\{1-\left[\rho_{12}\frac{\rho_{13}-\rho_{12}\rho_{23}}{1-\rho_{12}^2}+\rho_{23}\frac{\rho_{23}-\rho_{12}\rho_{13}}{1-\rho_{12}^2}\right]\right\}x_1x_2. \quad (18)$$

It is clear from (16), (17), and (18) that if $\mathbf{A}=(x_{ij})$, $i=1, 2, 3$; $j=1, 2, 3$; where the rows and columns of $\mathbf{A}_{11}$ correspond to the rows and columns of $\mathbf{P}_{11}$, then the joint distribution of x_1x_{3c} and x_2x_{3c}, given $\mathbf{A}_{11}$, is elliptical, $E(\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{A}_{11}, \mathbf{P}_{22.1}\oplus\mathbf{A}_{11}, \kappa_2)$.

(ii) Considering now the case of n independent sets of three variables each leading to a m.g.f. (14), it follows from the m.g.f. of the sums of squares and sums of products (after division by n), that the c.g.f. of the conditional distribution of

$$\frac{1}{n}\sum_{i=1}^{n}x_{1_i}x_{3_i}, \qquad \frac{1}{n}\sum_{i=1}^{n}x_{2_i}x_{3_i},$$

given

$$\frac{1}{n}\sum_{i=1}^{n}x_{1_i}^2, \frac{1}{n}\sum_{i=1}^{n}x_{1i}x_{2i}, \frac{1}{n}\sum_{i=1}^{n}x_{2i}^2,$$

is the sum of the c.g.f.'s of elliptical distributions with mean and variance given by (17) and (18). Thus the factors x_1^2, x_1x_2, and x_2^2 in (17) and (18) are replaced by the relevant mean products. The definition of the matrix $\mathbf{A}$ can thus be changed by replacing the elements x_ix_j with $(1/n)\sum_{r=1}^{n}x_{ir}x_{jr}$.

(iii) In the above consideration the moments about the population mean were used. However, the same reasoning as in the case of one kurtosis parameter holds also in the present situation. This shows that for moments about the sample means the parameter n should be equal to $N-1$, where N is the number of independent observations. Using now the notations $\mathbf{S}$, $\mathbf{S}_{11}$, and $\mathbf{S}_{22.1}$, as previously introduced, it follows from

$$\mathbf{S}_{22.1}\equiv s_{33c}=x_{33}-\frac{s_{22}s_{13}^2+s_{11}s_{23}^2-2s_{12}s_{13}s_{23}}{(s_{11}s_{22}-s_{12}^2)},$$

that setting $t_{33}=0$ in (16) did not remove the influence of s_{33} on relevant elements of the covariance matrix.

(iv) Return now to expression (14) for the general case, while noting that the conditional distribution $\mathbf{S}_{22.1}$ was already dealt with under (b) above, it follows that the values of all the generating variables t_{ij} associated with the elements in the rows and columns corresponding to $\mathbf{S}_{22}$ can be taken as zero. The procedure similar to the transition from (16) to (17) and (18)) becomes only a question of repeating the algebra in matrix notation when separating the marginal factor corresponding to $\mathbf{S}_{11}$ from the conditional factor corresponding to $\mathbf{S}_{21}$. Thus using the result in (i) of this section, it follows by replacing $\mathbf{A}_{11}$ by $\mathbf{S}_{11}$ that the general result can be stated as:

The conditional distribution of S_{21} *given* S_{11} *is* $E(\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{S}_{11}, \mathbf{P}_{22.1} \oplus \mathbf{S}_{11}, \kappa_2)$.

In this form the present result may be compared with the known result in multinormal theory given by Muirhead [3, p. 93].

4. An Application

As an application consider the following moments and correlation matrix, together with the conditional sampling correlations obtained from multitrait–multimethod data with respect to 50 persons measured for two traits (T) by two different methods (M) The four variables are X_1: ($T1M1$), X_2: ($T2M1$), X_3: ($T1M2$), X_4($T2M2$). The distribution of the conditional sampling matrix $\mathbf{S}_{22.1}$ given in (b) of Section 3, as well as the relations in (12), are taken as basis for discussion.

Moments	Mean	Var.	Skew.	Kur: γ_2	*Correlation*: X_1	X_2	X_3	X_4	*Cond. correlation*:
X_1	11.28	11.10	−0.04	−0.17	1.00	0.65	0.60	0.60	$r_{12.34} = 0.315$
X_2	10.20	9.06	0.13	−0.38	0.65	1.00	0.72	0.73	$r_{34.12} = 0.373$
X_3	15.62	64.36	−0.31	−1.60	0.60	0.72	1.00	0.72	
X_4	13.92	61.50	0.08	−1.53	0.60	0.73	0.72	1.00	

The relevant questions concern the degree to which the interdependence between traits (methods) can be attributed to method (trait). Although the underlying model to be discussed contains two kurtosis parameters, the conditional sampling matrix $\mathbf{S}_{22.1}$ depends on only one kurtosis parameter, so that using results in Muirhead [3, Chap. 5], the correlation structure can be treated in stead of the covariance structure.

Estimates of the two kurtosis parameters in (13) follow from equating the calculated marginal kurtosis parameters $\gamma_2/3$ to the expected kurtosis parameters of the marginal distributions of X_1, X_2, X_3, X_4 by using the relations similar to those given in [4, p. 9, Eq. (20]. The four expected kurtosis parameters (one-third of the relevant γ_2) follows directly from the relevant four coefficients of $t_1^4/8$, $t_2^4/8$, $t_3^4/8$, and $t_4^4/8$ in (13) in terms of the two parameters κ_1 and κ_2 and the population correlations. Substituting the calculated marginal kurtosis as well as the correlations as given in the above table, *four* linear moment equations in *two* population parameters κ_1 and κ_2 can be written down. A solution to the estimation problem in a similar case by combining the method of moments and of weighted least squares, was discussed in detail by the present author in [5]. However, in the present case, where the expected values of the coefficients of $t_1^4/8$ and $t_2^4/8$ are both equal to κ_1 an estimate $\hat{\kappa}_1 = -(0.17+0.38)/6 = -0.092$ seems feasible. Using this estimate of κ_1 in the two equations obtained from the coefficients of $t_3^4/8$ and $t_4^4/8$, two slightly different estimates -0.628 and -0.606 for κ_2 follow, giving an average estimate $\hat{\kappa}_2 = -0.617$. Amongst others, the conditional correlation between X_3 and X_4, that is, $r = r_{34.12}$ can now be transformed [3, p. 159] to $z = \frac{1}{2}\log[(1+r)/(1-r)]$ and considered as normal variate with variance $1/(n-2) + \kappa_2/(n+2)$, where $n+1$ is the sample size. Applying this result it follows that the value $r_{34.12} = 0.373$ in the above table is significant. A similar result is obtained on considering the correlation between X_1 and X_2, conditional on X_3 and X_4.

References

[1] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[2] Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya A* **32** 419–440.

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

[4] Steyn, H. S. (1993). On the problem of more than one kurtosis parameter in multivariate analysis. *J. Multivariate Anal.* **44** No. 1 1–22.

[5] Steyn, H. S. (1976). On the multivariate Poisson normal distribution. *J. Amer. Statis. Assoc.* **71**, No. 353 232–236.

[6] Sutradhar, B. C., and Ali, M. M. (1989). A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate *t* model. *J. Multivariate Anal.* **29** 155–162.