EXTRACTA MATHEMATICAE Vol. 8, Núm. 2–3, 98 – 101 (1993)

Multivariate Skewness and Kurtosis for Singular Distributions

R. ARDANUY AND J.M. SÁNCHEZ

Dpto. de Matemática Pura y Aplicada, Univ. de Salamanca,
Plaza de la Merced 1, E37008–*Salamanca, Spain*

AMS *Subject Class.* (1980): 62H10 *Received January* 18, 1993

INTRODUCTION AND NOTATIONS

In multivariate analysis it is generally assumed that the observations are normally distributed. It was Mardia ([1] to [5]), who first introduced measures of multivariate skewness and kurtosis; these statistics are affine invariant and can be used for testing multivariate normality. Skewness and kurtosis tests remain among the most powerful, general and easy to implement. In this paper we show some properties of these statistics when population distribution is singular.

Let $X=(X_1,X_2,\dots,X_p)^T$ be a random vector with mean vector $\mu=(\mu_1,\mu_2,\dots,\mu_p)^T$ and covariance matrix $\Sigma=(\sigma_{rs})$. When Σ is a regular matrix Mardia [1],[2] express the multivariate skewness, $\beta_1=\beta_{1p}$, and kurtosis, $\beta_2=\beta_{2p}$, as:

$$\beta_1=\beta_{1p}=\mathrm{E}[\{(X-\mu)^T\Sigma^{-1}(Y-\mu)\}^3] \tag{1}$$

$$\beta_2=\beta_{2p}=\mathrm{E}[\{(X-\mu)^T\Sigma^{-1}(X-\mu)\}^2] \tag{2}$$

being Y a random vector independent of X and identically distributed as X. Their sample counterpart are:

$$b_1=b_{1p}=\tfrac{1}{n^2}\textstyle\sum_{i,j=1}^{n}\{(X_i-\overline{X})^TS^{-1}(X_j-\overline{X})\}^3 \tag{3}$$

$$b_2=b_{2p}=\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\{(X_i-\overline{X})^TS^{-1}(X_j-\overline{X})\}^2 \tag{4}$$

being $X_i=(X_{1i},X_{2i},\dots,X_{pi})^T$, $i=1,2,\dots,n$, n independent observations on X, and $\overline{X}=(\overline{X}_1,\overline{X}_2,\dots,\overline{X}_p)^T$, and $S=(S_{rs})$ denote the sample mean vector and covariance matrix respectively. When Σ or S are singular matrices we may define the multivariate skewness and kurtosis using the g–inverse of Σ or S in the sense of Rao [7], i.e., $\Sigma\Sigma^{-}\Sigma=\Sigma$ and $SS^{-}S=S$:

(5) $$\beta_1 = \beta_{1p} = \mathrm{E}[\{(X-\mu)^T \Sigma^- (Y-\mu)\}^3]$$

(6) $$\beta_2 = \beta_{2p} = \mathrm{E}[\{(X-\mu)^T \Sigma^- (X-\mu)\}^2]$$

being Y a random vector independent of X and identically distributed as X. Their sample counterpart are:

(7) $$b_1 = b_{1p} = \frac{1}{n^2}\sum_{i,j=1}^{n}\{(X_i-\overline{X})^T S^- (X_j-\overline{X})\}^3$$

(8) $$b_2 = b_{2p} = \frac{1}{n}\sum_{i=1}^{n}\{(X_i-\overline{X})^T S^- (X_j-\overline{X})\}^2$$

MAIN RESULTS

PROPOSITION 1. *Expressions* (5) *to* (8) *remain the same for any* Σ^- *and any* S^-.

Proof. Since $X-\mu$ and $Y-\mu$ lies in the column space of Σ with probability 1, then there exist p–dimensional vectors U, V such that $X-\mu=\Sigma U$ and $Y-\mu=\Sigma V$, then:

$$(X-\mu)^T\Sigma^-(Y-\mu) = U^T\Sigma\Sigma^-\Sigma V = U^T\Sigma V$$
$$(X-\mu)^T\Sigma^-(X-\mu) = U^T\Sigma\Sigma^-\Sigma V = U^T\Sigma U$$

that not depend of g–inverse choosed. Analogously for the sample measures b_1 and b_2 we have that $X_i-\overline{X}$ lies in the column space of S, for i=1,2,..,n, then there exist p–dimensional vectors W_i such that $X_i-\overline{X}=\Sigma W_i$, $i=1,2,\ldots,n$, then:

$$(X_i-\overline{X})^T S^-(X_j-\overline{X}) = W_i^T SS^-SW_j = W_i^T SW_j$$

that also is independent of S^- for $i,j=1,2,\ldots,n$. This completes the proof. ∎

LEMMA 1. *Let* P,A *and* Q *matrices of appropiate sizes such that* $\mathrm{rank}(PAQ)=\mathrm{rank}(A)$, *then* $Q(PAQ)^-P$ *is a g–inverse of* A.

Proof. We have that $\mathrm{rank}(A)=\mathrm{rank}(PAQ)\leqslant\mathrm{rank}(PA)\leqslant\mathrm{rank}(A)=\mathrm{rank}(PAQ)\leqslant\mathrm{rank}(AQ)\leqslant\mathrm{rank}(A)$, hence $\mathrm{rank}(A)=\mathrm{rank}(PAQ)=\mathrm{rank}(PA)=\mathrm{rank}(AQ)$. Since $\mathrm{rank}(PAQ)=\mathrm{rank}(AQ)$ Corollary 1a.3 of Mitra [6] implies that $(PAQ)^-P$ is a g–inverse of AQ, also $\mathrm{rank}(AQ)=\mathrm{rank}(A)$ thus the same corollary implies that $Q(AQ)^-$ is a g–inverse of A. Therefore:

$$Q(PAQ)^-P = Q(AQ)^- = A^-$$

which concludes the proof. ∎

PROPOSITION 2. β_1, β_2, b_1 *and* b_2 *are invariant under those linear* *transformations which preserve the rank of covariance matrices* Σ *or* S *respectively.*

Proof. Let X and Y be independent random vectors with the same distribution with mean vector μ and covariance matrix Σ. Let $Z = AX + b$ be a linear transformation of random vector X such that $\text{rank}(\Sigma_Z) = \text{rank}(\Sigma)$. Since $\Sigma_Z = A\Sigma A^T$ and $\text{rank}(A\Sigma A^T) = \text{rank}(\Sigma)$ we have from Lemma 1 that $A^T(A\Sigma A^T)^- A$ is a g–inverse of Σ. Moreover $\mu_Z = A\mu + b$ and $W = AY + b$ is a random vector identically distributed as Z and independent of Z, thus:

$$\beta_1(Z) = \text{E}[\{(Z-\mu_Z)^T \Sigma_Z^- (W-\mu_Z)\}^3] = \text{E}[\{(X-\mu)^T A^T (A\Sigma A^T)^- A(Y-\mu)\}^3] = \\ = \text{E}[\{(X-\mu)^T \Sigma^- (Y-\mu)\}^3] = \beta_1(X)$$

and a similar proof implies that $\beta_2(Z) = \beta_2(X)$. For independent identically distributed observations $X_1, X_2, \ldots, X_n$, let $Z_i = AX_i + b$ be linear transformations of random vectors X_i, $i = 1, \ldots, n$, such that $\text{rank}(S_Z) = \text{rank}(S)$. Since $S_Z = ASA^T$ and $\text{rank}(ASA^T) = \text{rank}(S)$ we have from Lemma 1 that $A^T(ASA^T)^- A$ is a g–inverse of S. Moreover $\overline{Z} = A\overline{X} + b$ thus

$$b_1(Z) = \tfrac{1}{n^2} \textstyle\sum_{i,j=1}^n \{(Z_i - \overline{Z})^T S_Z^- (Z_j - \overline{Z})\}^3 = \\ = \tfrac{1}{n^2} \textstyle\sum_{i,j=1}^n \{(X_i - \overline{X})^T A^T (ASA^T)^- A(X_j - \overline{X})\}^3 = \\ = \tfrac{1}{n^2} \textstyle\sum_{i,j=1}^n \{(X_i - \overline{X})^T S^- (X_j - \overline{X})\}^3 = b_1(X)$$

and analogously we have that $b_2(Z) = b_2(X)$. ∎

PROPOSITION 3. *Let* $X_1, X_2, \ldots, X_n$ *be* n *independent observations from* $\text{N}_p(\mu, \Sigma)$ *such that* $n \geqslant r = \text{rank}(\Sigma)$, *then the sample distributions of* b_1 *and* b_2 *are the same than those obtained from a random sample* $Z_1, Z_2, \ldots, Z_n$ *from* $\text{N}_r(0, I_r)$.

Proof. Being Σ a symmetric nonnegative definite matrix of rank r, there exists an ortogonal matrix U such that:

$$U\Sigma U^T = \Lambda = \text{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0) = \begin{bmatrix} \Lambda_r & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & 0 \end{bmatrix}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_r$ are real and positive numbers. Consider the matrix $A = (\Lambda_r^{-1/2} \vdots 0)U$ and the linear transformation $Z = A(X - \mu)$, with X distributed as $\text{N}_p(\mu, \Sigma)$. Vector Z has a normal r–dimensional distribution with $\text{E}(Z) = 0$ and covariance matrix:

$$\Sigma_Z = A\Sigma A^T = (\Lambda_r^{-1/2} \vdots 0)UU^T\Lambda UU^T(\Lambda_r^{-1/2} \vdots 0)^T = (\Lambda_r^{-1/2} \vdots 0)\Lambda(\Lambda_r^{-1/2} \vdots 0)^T = I_r$$

then the vectors $Z_i = A(X_i - \mu)$, $i=1,\dots,n$, are a random sample from $N_r(0,I_r)$. Let S and S_Z be the sample variance matrices of X_i and Z_i, $i=1,\dots,n$, respectively. Since $n \geqslant r$ we have:

$$\operatorname{rank}(S_Z) = \operatorname{rank}(\Sigma_Z) = r = \operatorname{rank}(\Sigma) = \operatorname{rank}(S) \quad \text{a.s.}$$

then we can apply Proposition 2. Consequently $b_i(X) = b_i(Z)$, $i=1,2$. This completes the proof. ∎

REFERENCES

1. MARDIA, K.V., Measures of multivariate skewness and kurtosis with applications, *Biometrika* **57** (1970), 519–530.
2. MARDIA, K.V., Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, *Sankhyā: The Indian Journal of Statistics* **36**, Series B (1974), 115–128.
3. MARDIA, K.V., Assessment of multinormality and the robustness of Hotelling T^2 test, *Applied Statistics* **24** (1975), 163–171.
4. MARDIA, K.V., Algorithm AS 84: Measures of multivariate skewness and kurtosis, *Applied Statistics* **24** (1975), 262–265.
5. MARDIA, K.V., Tests of univariate and multivariate normality, in "Handbook of Statistics", Vol. 1, (P.R. Krishnaiah, Ed.), North–Holland Publishing Company, 1980, 279–320.
6. MITRA, S.J., On a generalised inverse of a matrix and applications, *Sankhyā: The Indian Journal of Statistics* **30**, Series A (1968), 107–114.
7. RAO, C.R., Calculus of generalised inverse of matrices. Part I: General theory, *Sankhyā: The Indian Journal of Statistics* **29**, Series A (1974), 317–342.