Standardized Moments of Bivariate Distributions

Anwar H Joarder
Department of Mathematical Sciences
King Fahd University of Petroleum and Minerals
Dhahran 31261 Saudi Arabia
Email: anwarj@kfupm.edu.sa

Abstract Theories have been developed to derive moments of the Mahalanobis distance of a bivariate distribution. The second and third order moments are related to kurtosis and skewness of the distribution. Ideas are illustrated with examples.

AMS Mathematics Subject Classification: 60E10, 60-02

Key Words and Phrases: bivariate distribution, standardized moments, Mahalanobis distance, product moments, kurtosis, skewness

1. Introduction

Product moments (also called raw product moments or product moments around zero) of order a and b for two random variables X_1 and X_2 are defined by $\mu'(a,b) = E\left(X_1^a X_2^b\right)$ while the centered product moments (sometimes called central product moments, corrected product moments or central mixed moments) are defined by

$$\mu(a,b) = E\left[\left(X_1 - E(X_1)\right)^a \left(X_2 - E(X_2)\right)^b\right]. \tag{1.1}$$

Interested readers may go through Johnson, Kotz and Kemp (1993, 46) or Johnson, Kotz and Balakrishnan (1997, 3). Evidently $\mu'(a,0) = E\left(X_1^a\right)$ is the a-th moment of X_1, and $\mu'(0,b) = E\left(X_2^b\right)$ is the b-th moment of X_2. In case X_1 and X_2 are independent, then $\mu'(a,b) = E\left(X_1^a\right)E\left(X_2^b\right) = \mu'(a,0)\mu'(0,b)$ and $\mu(a,b) = \mu(a,0)\mu(0,b)$. The correlation coefficient $\rho(-1 < \rho < 1)$ between X_1 and X_2 is denoted by

$$\rho_{X,Y} = \frac{\mu(1,1)}{\sqrt{\mu(2,0)\mu(0,2)}}. \tag{1.2}$$

Note that $\mu(2,0) = E\left(X_1 - E(X_1)\right)^2 = \sigma_{20}$ which is popularly denoted by σ_1^2 while the central product moment, $\mu(1,1) = E\left[\left(X_1 - E(X_1)\right)\left(X_2 - E(X_2)\right)\right]$ denoted popularly by σ_{12}, is, in fact, the covariance between X_1 and X_2.

In a series of papers, Mardia (1970, 1974, 1975) defined and discussed the properties of measures of kurtosis and skewness based on Mahalanobis distance. As it is difficult to derive the distribution of Mahalanobis distance for many distributions and calculate moments thereof, we derive Mahalanobis moments in terms of centered product moments. As an example Mahalanobis moments for a bivariate normal distribution is calculated. It is worth mentioning that the second Mahalanobis moment accounts for kurtosis while the third for skewness. Ideas are explained by examples.

In what follows we will rather use $X_1 = X$ and $X_2 = Y$ to avoid all confusion of a trivial nature, and define

$$\mu(a,b) = E\left[(X-\xi)^a (Y-\theta)^b\right]$$

where $\xi = E(X), \theta = E(Y)$.

2. Standardized Moments for any Bivariate Distribution

For a bivariate random vector $W = (X, Y)'$, with mean vector $\mu = (\xi, \theta)'$ and covariance matrix $Cov(W) = E(W-\mu)(W-\mu)' = \Sigma$,

the standardized distance is defined by

$$\begin{aligned} Q &= (W-\mu)'\Sigma^{-1}(W-\mu) \\ &= ((X-\xi)\ \ (Y-\theta))\Sigma^{-1}((X-\xi)\ \ (Y-\theta))' \end{aligned} \tag{2.1}$$

with

$$\Sigma = \begin{pmatrix} \mu(2,0) & \mu(1,1) \\ \mu(1,1) & \mu(0,2) \end{pmatrix}.$$

The quantity Q is also known to be generalized distance or Mahalanobis distance. The coefficients of kurtosis and skewness are $\beta_2 = E(Q^2)$ and $\beta_3 = E(Q^3)$ respectively (Kotz, Balakrishnan and Johnson, 2000, 77). For a bivariate random vector W with $E(W) = \mu$ and

$Cov(W) = \Sigma$, we define standardized moments by $\beta_i = E(Q^i),\ i = 1, 2, \cdots$ where $Q = (W - \mu)'\Sigma^{-1}(W - \mu) = \| \Sigma^{-1/2}(W - \mu) \|^2$.

As it is difficult to derive the distribution of Mahalanobis distance for many bivariate distributions and calculate moments thereof, we derive standardized moments in terms of centered product moments just to demonstrate the potential of an alternative way.

Theorem 2.1 Let X and Y have a bivariate distribution. Also let $\mu(a,b)$ be the centered product moment between X and Y, and $(\mu(2,0)\mu(0,2))^{-1/2}\mu(1,1) = \rho$ be the correlation coefficient between them. Then

$$(i)\ \mu(2,0)\mu(0,2)(1-\rho^2)E(Q) = 2\left[\mu(2,0)\mu(0,2) - \rho\mu^{1/2}(2,0)\mu^{1/2}(0,2)\mu(1,1)\right],$$

$$\begin{aligned}(ii)\ &\left[\mu(2,0)\mu(0,2)(1-\rho^2)\right]^2 E(Q^2)\\ &= \mu(4,0)\mu^2(0,2) + 4\rho^2\mu(2,0)\mu(0,2)\mu(2,2)\\ &+ \mu(0,4)\mu^2(2,0) - 4\rho\mu^{1/2}(2,0)\mu^{3/2}(0,2)\mu(3,1)\\ &+ 2\mu(2,0)\mu(0,2)\mu(2,2) - 4\rho\mu^{3/2}(2,0)\mu^{1/2}(0,2)\mu(1,3),\end{aligned}$$

$$\begin{aligned}(iii)\ &\left[\mu(2,0)\mu(0,2)(1-\rho^2)\right]^3 E(Q^3)\\ &= \mu(6,0)\mu^3(0,2) - 8\rho^3\mu^{3/2}(2,0)\mu^{3/2}(0,2)\mu(3,3) + \mu^3(2,0)\mu(0,6) - 6\rho\mu^{1/2}(2,0)\mu^{5/2}(2,0)\mu(5,1)\\ &+12\rho^2\mu(2,0)\mu^2(0,2)\mu(4,2) + 3\mu(2,0)\mu^2(0,2)\mu(4,2) + 3\mu^2(2,0)\mu(0,2)\mu(2,4)\\ &+12\rho^2\mu^2(2,0)\mu(0,2)\mu(2,4) - 6\rho\mu^{5/2}(2,0)\mu^{1/2}(0,2)\mu(1,5) - 12\rho\mu^{3/2}(2,0)\mu^{3/2}(0,2)\mu(3,3).\end{aligned}$$

Corollary 2.1 Let X and Y have a bivariate distribution with $E(X^aY^b) = E(X^bY^a)$ and the correlation between them is $\rho_{X,Y} = \rho$. Then

$$(i)\ (1-\rho^2)E(Q) = 2\left[1 - \rho\mu(1,1)\right],$$
$$(ii)\ \mu^2(0,2)(1-\rho^2)^2E(Q^2) = 2\mu(4,0) + (4\rho^2+2)\mu(2,2) - 8\rho\mu(3,1),$$
$$(iii)\ \mu^3(0,2)(1-\rho^2)^3E(Q^3) = 2\mu(6,0) - (8\rho^3+12\rho)\mu(3,3) - 12\rho\mu(5,1) + (24\rho^2+6)\mu(4,2).$$

Corollary 2.2 Let X and Y have a bivariate distribution. If X and Y are independent, then

$$(i)\ E(Q) = 2,$$

$(ii)\, E(Q^2) = 2 + \frac{\mu(4,0)}{\mu^2(2,0)} + \frac{\mu(0,4)}{\mu^2(0,2)},$

$(iii)\, E(Q^3) = \frac{\mu(6,0)}{\mu^3(2,0)} + \frac{\mu(0,6)}{\mu^3(0,2)} + 3\left(\frac{\mu(4,0)}{\mu^2(2,0)} + \frac{\mu(0,4)}{\mu^2(0,2)}\right).$

Corollary 2.3 Let X and Y have a bivariate distribution. If X and Y are independently and identically distributed, then

$(i)\; E(Q) = 2,$

$(ii)\, E(Q^2) = 2\left(1 + \frac{\mu(4,0)}{\mu^2(2,0)}\right),$

$(iii)\, E(Q^3) = 2\left(\frac{\mu(6,0)}{\mu^3(2,0)} + 3\frac{\mu(4,0)}{\mu^2(2,0)}\right).$

3. Product Moments of the Bivariate Normal Distribution

The pdf of the bivariate normal distribution is given by

$$f_1(x,y) = \frac{(1-\rho^2)^{-1/2}}{2\pi\sigma_1\sigma_2} \exp\left(\frac{-q(x,y)}{2}\right), \qquad (3.1)$$

where

$$(1-\rho^2)q(x,y) = \left(\frac{x-\xi}{\sigma_1}\right)^2 + \left(\frac{y-\theta}{\sigma_2}\right)^2 - \frac{2\rho(x-\xi)(y-\theta)}{\sigma_1\sigma_2}.$$

The following theorem is due to Kendal, Stuart (1969, 91).

Theorem 3.1 The centered product moments $\mu(a,b) = E\left[(X-\xi)^a (Y-\theta)^b\right]$ of the bivariate normal distribution with pdf in (3.1) are given by

$\mu(a,b) = \sigma_1^a \sigma_2^b \lambda(a,b)$ where

$\lambda(a,b) = (a+b-1)\rho\lambda(a-1,b-1) + (a-1)(b-1)(1-\rho^2)\lambda(a-2,b-2),$

$\lambda(2a,2b) = \frac{(2a)!(2b)!}{2^{a+b}} \sum_{j=0}^{\min(a,b)} \frac{(2\rho)^{2j}}{(a-j)!(b-j)!(2j)!},$

$\lambda(2a+1,2b+1) = \frac{(2a+1)!(2b+1)!}{2^{a+b}} \rho \sum_{j=0}^{\min(a,b)} \frac{(2\rho)^{2j}}{(a-j)!(b-j)!(2j+1)!},$

$\lambda(2a,2b+1) = \lambda(2a+1,2b) = 0.$

The above can be rewritten as

$\mu(a,b) = (a+b-1)\rho\sigma_1\sigma_2\mu(a-1,b-1) + (a-1)(b-1)(1-\rho^2)\sigma_1^2\sigma_2^2\mu(a-2,b-2),$

$$\mu(2a,2b)=\sigma_1^{2a}\sigma_2^{2b}\,\frac{(2a)!(2b)!}{2^{a+b}}\sum_{j=0}^{\min(a,b)}\frac{(2\rho)^{2j}}{(a-j)!(b-j)!(2j)!},$$

$$\mu(2a+1,2b+1)=\sigma_1^{2a+1}\sigma_2^{2b+1}\,\frac{(2a+1)!(2b+1)!}{2^{a+b}}\,\rho\sum_{j=0}^{\min(a,b)}\frac{(2\rho)^{2j}}{(a-j)!(b-j)!(2j+1)!},$$

$$\mu(2a,2b+1)=\mu(2a+1,2b)=0.$$

4. Standardized Moments of the Bivariate Normal Distribution

Let us represent the bivariate normal distribution with pdf in (3.1) by,

$W=\begin{pmatrix}X\\ Y\end{pmatrix}\sim N_2(\mu,\Sigma)$, $\mu=\begin{pmatrix}\xi\\ \theta\end{pmatrix}, \Sigma=\begin{pmatrix}\mu(2,0) & \mu(1,1)\\ \mu(1,1) & \mu(0,2)\end{pmatrix}$. It is known that for a p-variate normal distribution, $W\sim N_p(\mu,\Sigma)$, the standardized distance $Q=(W-\mu)'\Sigma^{-1}(W-\mu)\sim\chi_p^2$ so that $\beta_1=E(Q)=p$, $\beta_2=E(Q^2)=p(p+2)$ and $\beta_3=E(Q^3)=p(p+2)(p+4)$. That is for the univariate normal distribution, $\beta_1=1,\beta_2=3,\beta_3=15$ and for the bivariate normal distribution,

$$\beta_1=2,\beta_2=8,\beta_3=48. \tag{4.1}$$

Since the correlation between X and Y is ρ in the bivariate normal distribution, we apply the results of Section 3 to Theorem 2.1 to derive moments of standardized distance to demonstrate the potential of an alternative way.

Theorem 4.1 The first three standardized moments of the bivariate normal distribution are given by $\beta_1=2,\beta_2=8,\beta_3=48$.

Proof. It follows from Theorem 2.1 that

$$\begin{aligned}(i)\ &\sigma_1^2\sigma_2^2(1-\rho^2)E(Q)\\ &=2[\sigma_1^2\sigma_2^2-2\rho\sigma_1\sigma_2(\sigma_1\sigma_2\rho)]\\ &=2\sigma_1^2\sigma_2^2(1-\rho^2),\end{aligned}$$

$$\begin{aligned}(ii)&\left[\sigma_1^2\sigma_2^2(1-\rho^2)\right]^2E(Q^2)\\ &=3\sigma_1^4(\sigma_2^2)^2+4\rho^2\sigma_1^2\sigma_2^2[\sigma_1^2\sigma_2^2(1+2\rho^2)]\\ &+3\sigma_2^4(\sigma_1^2)^2-4\rho(\sigma_1^2)^{1/2}(\sigma_2^2)^{3/2}(3\sigma_1^3\sigma_2\rho)\\ &+2\sigma_1^2\sigma_2^2[\sigma_1^2\sigma_2^2(1+2\rho^2)]-4\rho(\sigma_1^2)^{3/2}(\sigma_2^2)^{1/2}(3\sigma_1\sigma_2^3\rho),\end{aligned}$$

so that

$$(1-\rho^2)^2 E(Q^2) = 3+4\rho^2(1+2\rho^2)+3-4\rho(3\rho)+2(1+2\rho^2)-4\rho(3\rho)$$
$$= 6+(4\rho^2+2)(1+2\rho^2)-24\rho^2$$
$$= 8(1-\rho^2)^2,$$

$$(iii)\left[\sigma_1^2\sigma_2^2(1-\rho^2)\right]^3 E(Q^3)$$
$$= 90\sigma_1^6(\sigma_2^2)^3 - 8\rho^3(\sigma_1^2)^{3/2}(\sigma_2^2)^{3/2}[3\sigma_1^3\sigma_2^3\rho(3+2\rho^2)]$$
$$+(\sigma_1^2)^3(90\sigma_2^6) - 6\rho(\sigma_1^2)^{1/2}(\sigma_2^2)^{5/2}[15\sigma_1^5\sigma_2\rho]$$
$$+12\rho^2\sigma_1^2(\sigma_2^2)^2[3\sigma_1^4\sigma_2^2(1+4\rho^2)]+3\sigma_1^2(\sigma_2^2)^2[3\sigma_1^4\sigma_2^2(1+4\rho^2)]$$
$$+3(\sigma_1^2)^2(\sigma_2^2)[3\sigma_1^2\sigma_2^4(1+4\rho^2)] + 12\rho^2(\sigma_2^2)^2(\sigma_2^2)[3\sigma_1^2\sigma_2^4(1+4\rho^2)]$$
$$-6\rho(\sigma_1^2)^{5/2}(\sigma_2^2)^{1/2}15\sigma_1\sigma_2^5\rho - 12\rho(\sigma_1^2)^{3/2}(\sigma_2^2)^{3/2}[3\sigma_1^3\sigma_2^3\rho(3+2\rho^2)]$$

so that

$$(1-\rho^2)^3 E(Q^3)$$
$$= 90-8\rho^3[3\rho(3+2\rho^2)]+(90)-6\rho[15\rho]+12\rho^2[3(1+4\rho^2)]$$
$$+3[3(1+4\rho^2)]+3[3(1+4\rho^2)] + 12\rho^2[3(1+4\rho^2)] - 6\rho(15\rho) - 12\rho[3\rho(3+2\rho^2)]$$
$$= 48(1-\rho^2)^3.$$

5. Standardized Moments for a Bivariate Chisquare Distribution

Joarder (2005) introduced the following chisquare distribution and checked that the distribution satisfies $E(U^aV^b) = E(U^bV^a)$, a property mentioned in Corollary 2.1.

Theorem 5.1 The random variables U and V are said to have a correlated bivariate chisquare distribution each with m degrees of freedom, if its probability density function is given by

$$f(u,v) = \frac{(uv)^{m/2-1}\, e^{-(u_1+u_2)}}{2^{m/2}\sqrt{\pi}\ \Gamma(m/2)} \sum_{k=0}^{\infty} \frac{(2uv)^{k/2}\,\Gamma\left((k+1)/2\right)}{k!\,\Gamma\left((k+m)/2\right)},$$

$m = N-1 > 2$.

Theorem 5.2 The (i, j) th moment of the distribution of U and V, is given by

$$E(U^iV^j) = \frac{2^{i+j}(1-\rho^2)^{i+j+m/2}}{\sqrt{\pi}\Gamma(m/2)} \sum_{k=0}^{\infty} \frac{(2\rho)^{k/2}}{k!}\,\Gamma\left(\frac{k+m}{2}+i\right)\Gamma\left(\frac{k+m}{2}+j\right)\frac{\Gamma\left((k+1)/2\right)}{\Gamma\left((k+m)/2\right)},$$

$m > 2\max(i, j)$.

Since $E(U^i) = 2^i\ \Gamma((m/2)+i)/\Gamma(m/2), i = 1, 2, \cdots$ we have $\mu(2,0) = E\left(U^2\right) - \xi^2 = 2m$, $\mu(1,1) = E(UV) - \xi^2 = m$, $\mu(0,2) = 2m$, $E(UV) = m\left(m+1\right)$, the correlation coefficient (see 1.2) between U and V is given by

$$\rho_{U,V} = \frac{E(U-E(U))(V-E(V))}{\sqrt{(E(U-E(U))^2(E(V-E(V))^2}} = \frac{\mu(1,1)}{\sqrt{\mu(2,0)\mu(0,2)}} = \frac{1}{2},$$

so that the covariance matrix between U and V is given by

$$\Omega = \begin{pmatrix} 2m & m \\ m & 2m \end{pmatrix}.$$

The standardized distance of the bivariate chisqaure distribution is given by $Q = W'\Omega^{-1}W$ where $W' = (U - m \;\; V - m)$ and the integer order standardized moments are $\beta_i = E\left(W'\Omega^{-1}W\right)^i$, $(i = 1, 2, \cdots)$. It is easy to check that $\beta_1 = E(W'\Omega^{-1}W) = 2$. Explicit expressions for the coefficients of kurtosis and skewness namely, β_2 and β_3 can be obtained by the use of Theorem 5.2.

Acknowledgement

The author gratefully acknowledges the excellent research support provided by King Fahd University of Petroleum and Minerals.

References

1. Anderson, T.W. (2003). An *Introduction to Multivariate Statistical Analysis*. John Wiley and Sons. New York.

2. Joarder, A.H. (2005) On a bivariate chisquare distribution. *Technical Report No. 335*, Dept of Mathematical Sciences, King Fahd University of Petroleum and Minerals, Saudi Arabia.

3. Johnson, N.L.; Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. John Wiley.

4. Johnson, N.L.; Kotz, S. and Kemp, A.W. (1993). *Univariate Discrete Distributions*. John Wiley.

5. Kendal, M.G. and Stuart, A. (1969). *The Advanced Theory of Statistics.* Charles Griffin and Co. Ltd, London.

5. Kotz, S.; Balakrishnan, N. and Johnson, N.L. (2000). *Continuous Multivariate Distributions*. John Wiley and Sons, New York.

6. Mardia, K.V. (1970a). A translation family of bivariate distributions and Frechet's bounds. *Sankhya*, 32A, 119-121.

7. Mardia, K.V. (1970b). *Families of Bivariate Distributions*. London: Griffin.

8. Mardia, K.V. (1970c). Measures of multivariate skewness and Kurtosis with applications. *Biometrika*, 57, 59-530.

9. Mardia, K.V. (1974). Applications of some measures of multivariate skewnewss and kurtosis

for testing normality and robustness studies. *Sankhya*, 36B, 115-128.

10. Mardia, K.V. (1975). Assessment of multinormality and the robustness of the Hotelling's T^2 test. *Applied Statistics*, 24, 163-171.

File: MSA\StandardizedMomentTRa.doc