Journal of Statistical Research *https://doi.org/10.3329/jsr.v59i1.83688*
2025, Vol. 59, No. 1, pp. 99-106 *ISSN 0256 - 422 X*

A NOTE ON THE SAMPLE COVARIANCE FROM A BIVARIATE NORMAL POPULATION

ANWAR H. JOARDER*
Department of Computer Science and Engineering
Northern University of Business and Technology Khulna, Khulna 9100, Bangladesh
Email: anwar.joarder@nubtkhulna.ac.bd

M. HAFIDZ OMAR
Department of Mathematics and Statistics
King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.
Email: omarmh@kfupm.edu.sa

SUMMARY

Both the Moment Generating Function and the Probability Density Function of sample covariance based on a bivariate normal distribution have been presented in a simpler way by using conditional distributions. We also prove that the random variable sample covariance is a linear combination of two chi-square random variables if they are statistically independent.

Keywords and phrases: Sample covariance; bivariate normal distribution; correlated chi-square variables; moment generating function

1 Introduction

Covariances have significant applications in finance and Modern Portfolio Theory (MPT). For example, in the capital asset pricing model (CAPM), which is used to calculate the expected return of an asset, the covariance between a security and the market is required.

Building on the work of Bose (1935), Mahalanobis et al. (1937) derived the probability density function (pdf) of the sample covariance based on a bivariate normal population. Bose derived the distribution of correlated variance ratio with the help of Fischer et al. (2023). It was shown by Pearson et al. (1929) that the sample covariance has the Bessel function distribution. An alternative proof is given by Kotz et al. (2001) in the Proposition 4.1.5. Press (1966) presented some characteristics of random sample covariance including the asymptotic distribution. Joarder and Omar (2011) also derived compact expressions for moments of sample covariance based on a bivariate normal population.

In Section 2, we overview the bivariate chi-square and a compounded normal distribution commonly known as the Variance Gamma (VG) Distribution. In Theorem 3.1, the Moment Generating

* Corresponding author

Function (MGF) of sample covariance is presented in a gentle way. In Theorem 3.2, we prove that the sample covariance is a linear combination of two chi-square random variables if and only if they are independent. In Section 4, we present a gentle derivation of the above century old problem of the probability density function of sample covariance for the modern readers of econometrics, business, computer science and others by using the VG Distribution.

2 Bivariate Chi-square and Compounded Normal Distribution

In this section, we overview the joint PDF of the two correlated chi-square random variables and a special compounded normal distribution better known as Variance Gamma (VG) Distribution.

Theorem 1. *(Joarder et al., 2012, Theorem 3.1) Let U and V be chi-square random variables based on a bivariate sample size $n = m + 1 \geq 2$ with correlation coefficient ρ^2 have the joint PDF (Probability Density Function) given by*

$$f(u, v) = \frac{(uv)^{(m-2)/2}}{2^m\, \Gamma^2(m/2)(1-\rho^2)^{m/2}} \exp\left(-\frac{u+v}{2-2\rho^2}\right) {}_0F_1\left(;\frac{m}{2};\frac{\rho^2 uv}{(2-2\rho^2)^2}\right), \tag{2.1}$$

$u > 0$, $v > 0$, with MGF (Moment Generating Function)

$$M_{U,V}(t_1, t_2) = [1 - 2(t_1 + t_2) + 4(1-\rho^2)t_1 t_2]^{-m/2}, \tag{2.2}$$

where $-1 < \rho < 1$ and ${}_0F_1(;b;z)$ is a generalized hypergeometric function defined in (A.2).

In case $\rho = 0$, the PDF in (2.1) would be that of the product of two independent chi-square random variables each having the same degrees of freedom m..

Let the gamma distribution $G(\alpha, \beta)$ have the shape parameter α and mean $\alpha\beta$ so that the MGF is $(1 - \beta t)^{-\alpha}$ where $\alpha > 0$, $\beta > 0$ and $t > 1/\beta$. If $X|V \sim N(\mu + \theta V, \sigma^2 V)$, $-\infty < x < \infty$, $-\infty < \theta < \infty, \sigma > 0, v > 0$, as assumed by Finaly and Seneta (2006), then $M_X(t) = E_V M_{X|V}(t)$ where $M_{X|V}(t) = \exp\{(\mu + \theta V)t + \sigma^2 V t^2/2\}$. Assuming $V \sim G(\alpha, 1/\alpha)$, $\alpha > 0$, the MGF of X simplifies to (2.4) with PDF in (2.3). The PDF in (2.3) is thus the compounded normal PDF with the mixture of mean $(\mu + \theta V)$and variance $(\sigma^2 V)$by a gamma random variable $V \sim G(\alpha, 1/\alpha)$, and can be derived by elementary probability distribution theory.

Theorem 2. *(Finaly and Seneta, 2006) If $X|V \sim N(\mu + \theta V, \sigma^2 V)$ and $V \sim G(\alpha, 1/\alpha)$, the PDF of $X$is given by*

$$f_X(x) = \sqrt{\frac{2}{\pi}} \frac{\alpha^\alpha}{\sigma\Gamma(\alpha)} \exp\left(\frac{\theta(x-\mu)}{\sigma^2}\right) \left(\frac{|x-\mu|}{\sqrt{\theta^2 + 2\alpha\sigma^2}}\right)^{\alpha - 1/2} K_{\alpha-1/2}\left(\frac{\sqrt{\theta^2 + 2\alpha\sigma^2}}{\sigma^2}|x - \mu|\right), \tag{2.3}$$

with MGF

$$M_X(t; \mu, \theta, \sigma, \alpha) = e^{\mu t}\left(1 - \frac{\theta t}{\alpha} - \frac{\sigma^2 t^2}{2\alpha}\right)^{-\alpha}, \tag{2.4}$$

where $-\infty < x < \infty$,$\alpha > 0$, $-\infty < \theta < \infty$,$\sigma > 0$,$-\infty < \mu < \infty$ and $K_\alpha(w)$ is defined by (A.3).

Proof. The PDF of $X|V \sim N(\mu+\theta V, \sigma^2 V)$ can be written as

$$f_{X|v}(x|v) = \frac{1}{\sigma^2\sqrt{2\pi V}} \exp\left[-\frac{1}{2}\frac{\{x-(\mu+\theta V)\}^2}{\sigma^2 V}\right].$$

Then the PDF in (2.3) of X can be derived by the following

$$f_X(x) = \int_{v=0}^{\infty} f_{X|v}(x|v) f_V(v) dv$$

where $f_V(v)$is the PDF of $V \sim G(\alpha, 1/\alpha)$.

The PDF in (2.3) is a special case of PDF (2.1) by Fischer et al. (2023). Following them, we will denote the PDF in (2.3) by $X \sim VG_2(\alpha, \theta, \sigma, \mu)$. See also equation (2.3) of Madan et al. (1998). If we put $\alpha = \lambda$,$\theta = 2b\lambda/(a^2-b^2)$, $\sigma^2 = 2\lambda/(a^2-b^2)$, in (2.3), we get the following hyperbolic form of the PDF of Variance Gamma Distribution

$$f_Y(y) = \frac{(a^2-b^2)^{\lambda} e^{b(y-\mu)}}{(2a)^{\lambda-1/2}\,\Gamma(\lambda)\sqrt{\pi}} |y-\mu|^{\lambda-1/2} K_{\lambda-1/2}(a|y-\mu|), -\infty < y < \infty, \tag{2.5}$$

where $\lambda > 0$, $-\infty < a < \infty$, $-\infty < b < \infty$, $-\infty < \mu < \infty$, $\sqrt{a^2-b^2} > 0$ with MGF

$$M_Y(t) = e^{\mu t}\left(\frac{a^2-b^2}{a^2-(b+t)^2}\right)^{\lambda},\ \lambda > 0,\ a^2 > (b+t)^2. \tag{2.6}$$

See for example, Gaunt (2014). □

3 The MGF of Sample Covariance and Some Related Characterizations

Let $\bar{X}$and $\bar{Y}$ be the means of two random variables $X$and Y based on a sample of size $n = m+1 \geq 2$. Then the sample covariance W is defined by

$$mW = \sum_{j=1}^{n} (X_j - \bar{X})(Y_j - \bar{Y}). \tag{3.1}$$

Let $X$and Ybe two random variables with the moment generating function $M_{X,Y}(s,t) = E(e^{sX+tY})$. Then for any t, $h < t < h$, $h > 0$, the random variables X and Y are independent if $M_{X,Y}(s,t) = M_X(t)\ M_Y(t)$ where $M_X(t) = E(e^{tX})$.

Let (X,Y) have a bivariate normal distribution (BND) with $E(X) = \mu_1$, $V(X) = \sigma_1^2$,$E(Y) = \mu_2$, $V(Y) = \sigma_2^2$ and $Cov(X,Y) = \rho\sigma_1\sigma_2$. Also let W be the sample covariance defined by (3.1) based on a sample of size $n = m+1$. For a BND, it can be easily proved that

$$(Y|X=x) \sim \eta\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\ (1-\rho^2)\sigma_2^2\right],$$

a conditional normal distribution. It can also be proved that $(mW|X_1 = x_1, X_2 = x_2, ..., X_n = x_n) \sim \eta\left[\rho\frac{\sigma_2}{\sigma_1}ms_1^2,\ (1-\rho^2)\sigma_2^2 ms_1^2\right]$, or,

$$(W|X_1 = x_1, X_2 = x_2, ..., X_n = x_n) \sim \eta\left[\rho\frac{\sigma_2}{\sigma_1}s_1^2,\ \frac{1}{m}(1-\rho^2)\sigma_2^2 s_1^2\right].$$

The following proof of the MGF of the sample covariance is known (Johnson, Kotz and Balakrishnan, 1995) and based on the conditional distribution of $W|X = \underset{\sim}{x}$ where $\underset{\sim}{x} = (X_1 = x_1, X_2 = x_2, ..., X_n = x_n)$. A brief proof is presented for broad spectrum of readers.

Theorem 3. *Let (X, Y) have a bivariate normal distribution with $E(X) = \mu_1$, $V(X) = \sigma_1^2$, $E(Y) = \mu_2$, $V(Y) = \sigma_2^2$ and $Cov(X, Y) = \rho\sigma_1\sigma_2$ where $\sigma_1 > 0$, $\sigma_2 > 0$ and $1 \le \rho \le 1$. Also, let the sample covariance be defined by* (3.1). *Then the MGF of W defined by* (3.1) *is given by Equation **(a)**:*

$$M_W(t) = \left[1 - 2\rho\left(\frac{\sigma_1\sigma_2}{m}\right)t - (1-\rho^2)\left(\frac{\sigma_1\sigma_2}{m}\right)^2 t^2\right]^{-m/2} \tag{3.2}$$

*or by Equation **(b)**:*

$$M_W(t) = [(1-2bt)(1-2ct)]^{-m/2} \tag{3.3}$$

where b and c are defined by $2mb = \sigma_1\sigma_2(1+\rho)$ and $2mc = -\sigma_1\sigma_2(1-\rho)$ respectively.

Proof. **Proof of Part (a):** We know that the moment generating function of $X \sim N(\mu, \sigma^2)$ is given by

$$M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right), \quad t \in R.$$

Then for the bivariate normal distribution, we have

$$(W|\underset{\sim}{X} = \underset{\sim}{x}) \sim \eta\left[\rho\frac{\sigma_2}{\sigma_1}s_1^2,\ \frac{1}{m}(1-\rho^2)\sigma_2^2 s_1^2\right].$$

It has the following MGF:

$$M_{W|\underset{\sim}{x}}(t) = E\left(e^{tW}|\underset{\sim}{X} = \underset{\sim}{x}\right) = \exp\left[\rho\frac{\sigma_2}{\sigma_1}s_1^2 t + \frac{1}{2}\ t^2\frac{1}{m}(1-\rho^2)\sigma_2^2 s_1^2\right],$$

which can be simplified to be

$$E\left(e^{tW}|\underset{\sim}{X} = \underset{\sim}{x}\right) = \exp\left[\left\{\frac{\rho\sigma_1\sigma_2}{m}t + \ \frac{1-\rho^2}{2}\left(\frac{\sigma_1\sigma_2}{m}\right)^2 t^2\right\}\frac{ms_1^2}{\sigma_1^2}\right].$$

Letting $d = \frac{\rho\sigma_1\sigma_2}{m}t + \ \frac{1-\rho^2}{2}\left(\frac{\sigma_1\sigma_2}{m}\right)^2 t^2$ and $u = \frac{ms_1^2}{\sigma_1^2}$, we have

$$E\left(e^{tW}|\underset{\sim}{X} = \underset{\sim}{x}\right) = e^{du}.$$

Then we have

$$M_{W|\underset{\sim}{X}}(t) = E\left(e^{tW}|\underset{\sim}{X}\right) = e^{dU}, \quad where \;\; U = \frac{mS_1^2}{\sigma_1^2} \sim \chi_m^2.$$

Now taking expectation over $\underset{\sim}{X}$, we have $E\left[E\left(e^{tW}|\underset{\sim}{X}\right)\right] = M_U(d)$, the MGF of $U \sim \chi_m^2$ so that

$$E\left[E\left(e^{tW}|\underset{\sim}{X}\right)\right] = (1-2d)^{-m/2}.$$

But the left hand side is the $E\left(e^{tW}\right) = M_W(t)$, i.e.,

$$M_W(t) = \left[1 - 2\left\{\frac{\rho\sigma_1\sigma_2}{m}t + \frac{1-\rho^2}{2}\left(\frac{\sigma_1\sigma_2}{m}\right)^2 t^2\right\}\right]^{-m/2}$$

which is the same as (3.2) or (32.123) of Johnson et.al. (1995).

Proof of Part (b): The inner quadratic trinomial of MGF of W in (3.2), say, $q(t)$, can be written as

$$q(t) = 1 - [(1+\rho) - (1-\rho)]\left(\frac{\sigma_1\sigma_2}{m}\right)t - (1-\rho^2)\left(\frac{\sigma_1\sigma_2}{m}\right)^2 t^2,$$
$$q(t) = 1 - 2bt - 2ct + 4bct^2.$$

The above can be factored as $q(t) = (1-2bt)(1-2ct)$. The proof is thus complete. □

The presentation of the following theorem in Johnson, Kotz and Balakrishnan (1995) is improved below.

Theorem 4. *Let (X,Y) have a bivariate normal distribution with $E(X) = \mu_1, V(X) = \sigma_1^2, E(Y) = \mu_2, V(Y) = \sigma_2^2$ and $Cov(X,Y) = \rho\sigma_1\sigma_2$. Also let the sample covariance W be defined by* (3.1) *based on a sample of size $n = m+1 \geq 2$. Also let b and c be defined by $2mb = \sigma_1\sigma_2(1+\rho)$ and $2mc = -\sigma_1\sigma_2(1-\rho)$ where each of the random variables $U$and V be defined by χ_m^2, a chi-square random variable with m degrees of freedom. Then the random sample covariance can be expressed as $W = bU + cV$ if and only if U and V are statistically independent.*

Proof. (**Proof of the First Part:**) By definition, the MGF of $bU + cV$is given by

$$M_{bU+cV}(t) = E\left[e^{t(bU+cV)}\right] = E\left[e^{(tb)U}\ e^{(tc)V}\right].$$

By statistical independence of $U$and V, we have

$$M_{bU+cV}(t) = E\left[e^{(tb)U}\right]E\left[e^{(tc)V}\right], \quad or, \quad M_{bU+cV}(t) = M_U(bt)\ M_V(ct).$$

Since $U \sim \chi_m^2$and $V \sim \chi_m^2$, we have $M_{bU+cV}(t) = (1-2bt)^{-m/2}\ (1-2ct)^{-m/2}$, which is the MGF of $bU + cV$. But the above is also the MGF of W. By the uniqueness theorem of MGF, if $U$and $V$are independent, then Whas the characterization $W = bU + cV$.

(**Proof of the Second Part:**) For any two random variables $U \sim \chi_m^2$ and $V \sim \chi_m^2$, the MGF of the quantity $bU + cV$is given by

$$M_{bU+cV}(t) = E\left[e^{t(bU+cV)}\right] = E\left(e^{tbU}\ e^{tcV}\right). \tag{3.4}$$

The right hand side of the MGF of W in (3.3) can be written as

$$(1-2bt)^{-m/2}(1-2ct)^{-m/2} = E\left[e^{(tb)U}\right]\left[e^{(tc)V}\right], \tag{3.5}$$

where $U \sim \chi_m^2$ and $V \sim \chi_m^2$. Obviously, the MGF of $bU + cV$ in (3.4) and that of W in (3.3) or equivalently in (3.5) will be the same, i.e., if

$$E\left(e^{tbU}\ e^{tcV}\right) = E\left[e^{(tb)U}\right]\left[e^{(tc)V}\right],$$

i.e., if the random variables U and V are independent. By the uniqueness theorem of MGF, we have proved that $W = bU + cV$ if the random variables U and V are independent. □

4 The PDF of Sample Covariance by Variance Gamma Distribution

The following derivation is briefly mentioned in Fischer et al. (2023).

Theorem 5. *Let W be the sample covariance defined by* (3.1) *based on a sample $n = m + 1 \geq 2$ from a bivariate normal distribution. Then the PDF of W is given by*

$$f_W(w) = C_{\sigma_1,\sigma_2}^{m,\rho}|w|^{(m-1)/2}\exp\left(\frac{\rho m w}{(1-\rho^2)\sigma_1\sigma_2}\right)\ K_{(m-1)/2}\left(\frac{m|w|}{(1-\rho^2)\sigma_1\sigma_2}\right), \tag{4.1}$$

$$-\infty < w < \infty, 2^{(m-1)/2}\sqrt{\pi(1-\rho^2)}\Gamma\left(m/2\right)C_{\sigma_1,\sigma_2}^{m,\rho} = (m/\sigma_1\sigma_2)^{(m+1)/2}, \sigma_1 > 0,\ \sigma_2 > 0,$$

$-1 < \rho < 1$ *and $K_\alpha(x)$ is the modified Bessel function of the second kind represented by (A.3).*

Proof. If $\alpha = m/2$, $\theta = \rho\sigma_1\sigma_2$, $\sigma^2 = (1-\rho^2)(\sigma_1\sigma_2)^2/m$, $\mu = 0$, in (2.4), we get the MGF of W in (3.2) and then by the pdf in (2.3) we get the pdf in (4.1) which will be denoted by $W \sim VG_2\left(\alpha = m/2,\ \theta = \rho\sigma_1\sigma_2,\ \sigma^2 = (1-\rho^2)(\sigma_1\sigma_2)^2/m,\ \mu = 0\ \right)$. The pdf in (4.1) can also be derived by (2.6) and (2.5).

□

Interested readers can go through moments and related characteristics derived by Joarder and Omar (2011). The graph of the PDF of covariance is drawn in Appendix 2 for several values of ρ.

5 Conclusion

Though there have been huge research on sample correlation coefficient for over a century, we do not see much on the sample covariance. Recently sample covariance has been applied to many business investigations. We hope the research will lead to further investigation in mathematical statistics and find applications in business, econometrics, finance etc.

Declarations Conflict of interest

The authors declare that they have no conflict of interest.

References

Finaly, R., and Seneta, E. (2006), "Stationary-increment Student and variance-gamma processes," *Journal of Applied Probability*, 43, 441–453.

Fischer, A., Gaunt, R. E., and Sarantsev, A. (2023), "The variance gamma distribution: A review," available at `https://arxiv.org/pdf/2303.05615.pdf`.

Bose, S. S. (1935), "On the distribution of the ratio of two variances of two samples drawn from a given bivariate correlated population," *Sankhya*, 2(1), 65–72.

Fisher, R. A. (1915), "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," *Biometrika*, 10, 507–521.

Gaunt, R. E. (2014), "Variance-Gamma approximation by Stein's method," *Electronic Journal of Probability*, 19(38), 1–33.

Joarder, A. H., and Omar, M. H. (2011), "On the moments of sample covariance from a bivariate normal population," *Journal of Applied Statistical Science*, 19(3), 261–272.

Joarder, A. H., Laradji, A., and Omar, M. H. (2012), "On some characteristics of bivariate chi-square distribution," *Statistics*, 46(5), 577–586.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995), *Continuous Univariate Distributions, Vol. 2*, New York: John Wiley and Sons.

Mahalanobis, P. C., Chandra, R., and Roy, S. N. (1937), "Normalisation of statistical variates and the use of rectangular co-ordinates in the theory of sampling distributions," *Sankhya*, 3(1), 1–40.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998), "The variance gamma processes and option pricing," *European Finance Review*, 2, 79–105.

Pearson, K., Jefferey, G. B., and Elderton, E. M. (1929), "On the distribution of the first product moment-coefficient in samples drawn from an indefinitely large normal population," *Biometrika*, 21, 164–193.

Seneta, E. (2004), "Fitting the variance-gamma model to financial data," *Journal of Applied Probability: Stochastic Methods and Their Applications*, 41, 177–187.

A Appendix 1

We will be using the following product of $k$consecutive ascending integers:

$$a_{\{k\}} = a(a+1)\cdots(a+k-1), a_{\{0\}} = 1 \tag{A.1}$$

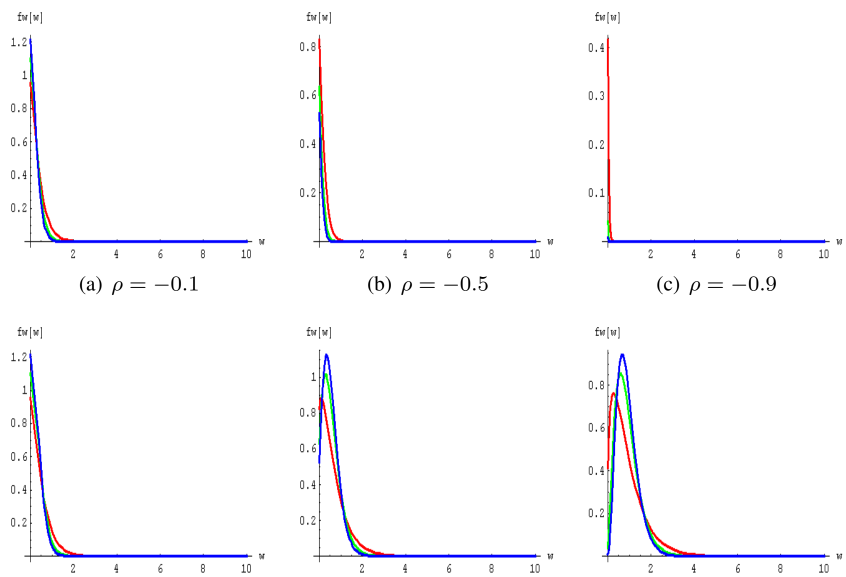
The generalized hypergeometric function ${}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots, b_q; z)$ is defined by

$${}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots, b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_{\{k\}}(a_2)_{\{k\}} \cdots (a_p)_{\{k\}}}{(b_1)_{\{k\}}(b_2)_{\{k\}} \cdots (b_q)_{\{k\}}} \frac{z^k}{k!}, \tag{A.2}$$

where $a_{\{k\}}$ is defined in (A.1). The modified Bessel function of the second kind admits numerous integral representations. We present the following (Seneta, 2004):

$$K_\alpha(w) = \frac{1}{2}\int_0^\infty x^{\alpha-1} \exp\left[-\frac{w}{2}\left(x + \frac{1}{x}\right)\right] dx,\ w > 0. \tag{A.3}$$

B Some graphs of pdf of sample covariance



(a) $\rho = -0.1$ (b) $\rho = -0.5$ (c) $\rho = -0.9$

(d) $\rho = 0.1$ (e) $\rho = 0.5$ (f) $\rho = 0.9$

Received: March 7, 2025

Accepted: June 18, 2025