King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 432

Feb 2014

A New Method for the Covariance Between Sample Mean and Variance

Anwar H. Joarder and A. Laradji

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: c-math@kfupm.edu.sa

A New Method for the Covariance Between Sample Mean and Variance

Anwar H. Joarder
Mathematical and Computing Sciences
Universiti Brunei Darussalam
Gadong BE1410, Brunei Darusslam
Email: ajstat@gmail.com

A. Laradji
Department of Mathematics and Statistics
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia
Email: alaradji@kfupm.edu.sa

Abstract. We derive a new expression for the covariance between sample mean and variance in terms of the moments of sample observations. Several corollaries are deduced. We also derive a general expression for the covariance between linear and quadratic functions of sample observations.

Key Words: Covariance, sample mean, sample variance, independence

MSC (2010). 62E10, 62H20, 62G35.

1. Introduction

The independence of sample mean and variance facilitates the derivation of Student t-statistic. The probability density function of the t-statistic was derived by Student (1908) under the assumption of the normality of the parent population. While he used uncorrelation of sample mean and variance, it was Fisher (1925) who clearly mentioned that the independence between sample mean and variance simplifies the derivation and thus defined the t-statistic as we define today. The confidence interval for the mean of a normal population with unknown variance, and the test of significance of the mean, testing equality of two means and a huge body of statistical methods are based on the t-statistic.

Geary (1936) is the first to prove that the sample mean and variance are independent if and only if the sample observations are independently, identically and normally distributed. Recently, Mukhopadhyay (2005) came up with some examples that shows that the condition of normality is not a necessity. However, the sample mean and variance are not generally independent in wider class of distributions. Hence it is desirable to know the covariance or correlation between sample mean and variance.

It is known that the covariance between sample mean and variance for independently and identically distributed random variable is proportional to the third central moment. See for example, Kang (1984). Kang and Goldsman (1990) exhibited the uncorrelated behavior of sample mean and variance by computer simulation.

Recently, Zhang (2007) addressed the issue. Mukhopadhah and Son (2011) derived the covariance between sample mean $\overline{X}$ and variance $S^2(X)$ based on correlated and non-identically distributed observations in terms of covariance of transformed variables $Y_j = X_j - \mu_j \ \ (j = 1, 2, \ldots, n)$. They also have provided insightful examples.

In Section 3, we utilize a general expression (3.1) to derive the covariance between sample mean and variance. It has been specialized to some known and new cases. For variables that are identically but not independently distributed, we have got a simple expression for covariance between sample mean and variance and presented in Corollary 3.2. Finally, in Section 4, we derive a new expression for the covariance between linear and quadratic functions of sample observations.

2. Some Preliminaries

Let $X_1, X_2, \ldots, X_n \ \ (n = 2, 3, \cdots)$ have an arbitrary n-dimensional joint distribution. We define the sample mean $\overline{x}$ and variance s^2 by $n\overline{x} = \sum_{i=1}^{i=n} x_i$ and $(n-1)s^2 = \sum_{i=1}^{i=n} (x_i - \overline{x})^2$, respectively. The sample variance can also be represented by $n(n-1)s^2 = (n-1)\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} \sum_{j(\neq i)=1}^{n} x_i x_j$, or, by

$$2n(n-1)S^2 = \sum_{i=1}^{i=n} \sum_{j(\neq i)=1}^{j=n} (X_i - X_j)^2. \qquad (2.1)$$

Also for identically distributed observations $X_1, X_2, \ldots, X_n$ with common mean μ, we denote $\mu'_a \equiv E(X^a)$, the a-th moment of X and $\mu_a \equiv E(X-\mu)^a$, the a-th centered moment of X. order Then the mean μ'_1 and variance $\mu_2 \equiv V(X)$ will be simply denoted by μ and $\sigma^2 = \mu'_2 - \mu^2$ respectively.

3. Covariance Between Sample Mean and Variance: A Simpler Approach

Let $X_1, X_2, \cdots, X_n \ \ (n = 2, 3, \cdots)$ have an arbitrary n-dimensional joint distribution. Let

$$\gamma_{kij} = Cov\left[X_k, (X_i - X_j)^2\right], \quad i = 1, 2, \ldots, n; \ j(\neq i) = 1, 2, \ldots, n. \qquad (3.1)$$

Then γ_{kij} can be explicitly written as

$$\gamma_{kij} = E(X_k X_i^2) - 2E(X_k X_i X_j) + E(X_k X_j^2) - E(X_k)E(X_i^2) + 2E(X_k)E(X_i X_j) - E(X_k)E(X_j^2).$$

The following lemma is obvious.

Lemma 3.1 Let $X_1, X_2, \cdots, X_n \ \ (n = 2, 3, \cdots)$ be independently and identically distributed. Then
$Cov\left[X_1, \ (X_1 - X_2)^a \ \right] = \sigma^2, \ \mu_3$ and $\mu_4 + 3\sigma^4$ for $a = 1, 2, 3.$ (3.2)

Theorem 3.1 Let $X_1, X_2, \cdots, X_n$ $(n \geq 2)$ have an arbitrary n-dimensional joint distribution.

a. If the variables are independently but not identically distributed , then

$$\gamma_{iij} = E(X_i^3) - E^3(X_i) - E(X_i)\sigma_i^2 - 2E(X_j)\sigma_i^2, \ \ i = 1, 2, \cdots, n; \ j(\neq i) = 1, 2, \cdots, n. \tag{3.3}$$

b. If the variables are independently but not identically distributed and $1 \leq k \leq n, (k \neq i),$ then the following holds:

$$\sum_{i=1}^{n} \sum_{j(\neq i)=1}^{j=n} \gamma_{kij} = 2\sum_{j(\neq k)=1}^{n} \gamma_{kkj}. \tag{3.4}$$

c. If the variables are identically but not independently distributed and $1 \leq k \leq n,$ then

$$\sum_{i=1}^{n} \sum_{j(\neq i)=1}^{j=n} \gamma_{kij} = 2(n-1)\left[\mu_3' - \mu_{21}' - 2\mu(1-\rho)\sigma^2\right] + (n-1)(n-2)\gamma_{123}, \tag{3.5}$$

where γ_{123} is defined in (3.1).

Proof. Proof of part (a) is straightforward, and part (b) follows from the fact that X_k and $(X_i - X_j)^2$ are independent when k, i and j are distinct.

c. Since the variables are identically distributed, we can assume without loss of generality that $k = 1$. We have

$$\sum_{i=1}^{n} \sum_{j(\neq i)=1}^{j=n} \gamma_{1ij} = \sum_{i=1}^{n} \sum_{j(\neq i)=1}^{j=n} Cov\left[X_k, \ (X_i - X_j)^2\right]. \tag{3.6}$$

Expanding the above sum, and by noting that the variables are identically but not independently distributed, it is easy to see that $2(n-1)$ of its terms are equal to $Cov\left(X_1, (X_1 - X_2)^2\right) = \gamma_{112}$ and the remaining $(n-1)(n-2)$ terms are equal to $\operatorname{cov}\left(X_1, (X_2 - X_3)^2\right) = \gamma_{123}$. Also, $\gamma_{112} = E(X_1^{\ 3}) - E(X_1^{\ 2} X_2) - 2E(X_1)\left[E(X_1^{\ 2}) - E(X_1 X_2)\right]$ which simplifies to $\mu_3' - \mu_{21}' - 2\mu\sigma^2(1-\rho)$. This proves (3.5).

Theorem 3.2 Let $X_1, X_2, \cdots, X_n$ $(n = 2, 3, \cdots)$ have an arbitrary n-dimensional joint distribution. Then the covariance between $\bar{X}$ and S^2 is given by

a. $n(n-1)Cov(\bar{X}, \ S^2) = \sum_{k=1}^{n} Cov\left[X_k, \ (n-1)S^2\right],$ (3.7)

where

$$Cov\left[X_k, \ (n-1)S^2\right] = Cov(X_k, \ X_k^2) + \sum_{i(\neq k)=1}^{n} Cov(X_k, \ X_i^2) - \frac{1}{n} Cov\left[X_k, (n\bar{X})^2\right]. \tag{3.8}$$

b. $2n^2(n-1)Cov(\bar{X}, \ S^2) = \sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{j(\neq i)=1}^{j=n} Cov\left[X_k, \ (X_i - X_j)^2\right]$ (3.9)

Proof.

a. The expression in (3.7) is obvious. By expanding S^2, we have

$$Cov\left[X_k,\ (n-1)S^2\right] = E\left[X_k(X_1^2+X_2^2+\cdots+X_n^2)\right] - \frac{1}{n}E\left[X_k(n\bar{X})^2\right]$$
$$-\left[E(X_k)E(X_1^2+X_2^2+\cdots+X_n^2) - \frac{1}{n}E(X_k)E(n\bar{X})^{22}\right].$$

The above simplifies to (3.8).

b. It is easy to see that $nCov(\bar{X},S^2) = \sum_{k=1}^{n} Cov(X_k,S^2)$. Since the sample variance S^2 has the representation $2n(n-1)S^2 = \sum_{i=1}^{i=n}\sum_{j(\neq i)=1}^{j=n}(X_i - X_j)^2$, we have

$Cov(X_k,S^2) = E(X_kS^2) - E(X_k)E(S^2)$, or,

$$Cov(X_k,S^2) = \frac{1}{2n(n-1)}Cov\left[X_k,\ \sum_{i=1}^{i=n}\sum_{j(\neq i)=1}^{j=n}(X_i-X_j)^2\right]. \tag{3.10}$$

Then we have

$$2n^2(n-1)Cov(\bar{X},\ S^2) = \sum_{k=1}^{n} Cov\left[X_k,\ \sum_{i=1}^{n}\sum_{j(\neq i)=1}^{j=n}(X_i-X_j)^2\right], \tag{3.11}$$

which is equivalent to (3.9).

If $X_1, X_2, \cdots, X_n$ are independently distributed, then from (4.8), we have

$$Cov\left[X_k,\ (n-1)S^2\right] = Cov(X_k, X_k^2) - \frac{1}{n}Cov\left[X_k, (n\bar{X})^2\right].$$

The following corollary is due to Mukhopaddahay and Sun (2011). An alternative proof is presented here.

Corollary 3.1 Let $X_1, X_2, \cdots, X_n$ be independently but not identically distributed. Also let $E(X_i)$ and $\sigma_i^2 = E[X_i - E(X_i)]^2$, be the mean and variance of any X_i, $i = 1,2,\cdots,n.$ Then

$$Cov(\bar{X},\ S^2) = \frac{1}{n^2}\left[\sum_{i=1}^{i=n}\left(E(X_i^3) - [E(X_i)]^3\right)\right] - \frac{n-3}{n^2(n-1)}\sum_{i=1}^{i=n}E(X_i)\sigma_i^2$$
$$-\frac{2}{n^2(n-1)}\left(\sum_{i=1}^{i=n}E(X_i)\right)\left(\sum_{i=1}^{i=n}\sigma_i^2\right), \tag{3.12}$$

for $n = 2,3,\cdots.$

Proof. Using (3.4) in (3.9), we have

$$n^2(n-1)Cov(\bar{X},\ S^2)=\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\gamma_{iij}. \tag{3.13}$$

Since the variables are independently but not identically distributed, by part (a) of Theorem 3.1, we have $\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\gamma_{iij}=\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\left[E(X_i^3)-\nu_i^3-\nu_i\sigma_i^2-2\nu_j\sigma_i^2\right]$ where $\nu_i=E(X_i),\ i=1,2,\ldots,n.$

The above can be written as

$$\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\gamma_{iij}=(n-1)\sum_{i=1}^{n}\left[E(X_i^3)-\nu_i^3\right]-(n-1)\sum_{i=1}^{n}\nu_i\sigma_i^2+2\sum_{i=1}^{n}\nu_i\sigma_i^2-2\left(\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\nu_j\sigma_i^2+\sum_{i=1}^{n}\nu_i\sigma_i^2\right).$$

The proof is complete by virtue of $\left(\sum_{j=1}^{n}\nu_j\right)\left(\sum_{i=1}^{n}\sigma_i^2\right)=\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\nu_j\sigma_i^2+\sum_{i=1}^{n}\nu_i\sigma_i^2.$

In case $n=3,$ we have

$$9Cov(\bar{X},\ S^2)=\left[\sum_{i=1}^{i=3}\left(E(X_i^3)-[E(X_i)]^3\right)\right]-\left(\sum_{i=1}^{i=3}E(X_i)\right)\left(\sum_{i=1}^{i=3}\sigma_i^2\right).$$

We remark that Corollary 3.1 matches with Corollary 2.1 of Mukhopadhay and Son (2011) for $Y_i=X_i-E(X_i),\ i=1,2,\cdots,n.$ Furthermore, if $X_1,X_2,\cdots,X_n$ are independently and normally distributed, then

$E(X_i^3)=E(X_i)\left[\left(E(X_i)\right)^2+3\sigma_i^2\right],$ and

$$Cov(\bar{X},\ S^2)=\frac{2}{n(n-1)}\left[\sum_{i=1}^{i=n}E(X_i)\sigma_i^2-\frac{1}{n}\left(\sum_{i=1}^{i=n}E(X_i)\right)\left(\sum_{i=1}^{i=n}\sigma_i^2\right)\right],$$

as in Example 2.2 of Mukhopadhay and Son (2011). They demonstrated that the covariance may be negative, zero or positive depending on the choices of $E(X_i)$ and $\sigma_i\ \ (i=1,2,3).$

Corollary 3.2 Let $X_1,X_2,\cdots,X_n$ be identically but not independently distributed. Then

$$2nCov(\bar{X},\ S^2)=2[\mu_3'-\mu_{21}'-2\mu(1-\rho)\sigma^2]+(n-2)\gamma_{123}, \tag{3.14}$$

where $\gamma_{123}=Cov\left[X_1,\ (X_2-X_3)^2\right].$

Proof. Use (3.5) in (3.9).

Corollary 3.3 Let $X_1,X_2,\cdots,X_n$ be independently and identically distributed. Then

$$Cov(\bar{X},S^2)=\mu_3/n. \tag{3.15}$$

Proof. In case the variables are independently and identically distributed, then $\rho=0$ and $\gamma_{123}=0$ and then from (3.14) we have $nCov(\bar{X},\ S^2)=\mu_3.$ Also by using $E(X_i)=\mu,\ \ \sigma_i^2=\sigma^2$ for any $i=1,2,\cdots,n,$ in (3.12), we have $nCov(\bar{X},\ S^2)=\mu_3.$

Example 3.1 Let $X_1, X_2, \cdots X_n$ be independently and identically distributed Bernoulli random variables $B(1, p),\ 0 < p < 1.$ (Zhang, 2007).

Let $Y = X_1 + X_2 + \cdots + X_n$ so that $\bar{X} = Y/n$ and $S^2 = Y(n-Y)/n(n-1).$ Since $Y \sim B(n, p),$ a binomial random variable, $n^2 Cov(\bar{X},\ S^2) = Cov[Y, Y(n-Y)],$ or, $Cov(\bar{X},\ S^2) = \mu_3/n,$ where $\mu_3 = p(1-p)(1-2p),$ the third central moment of Bernoulli distribution. Clearly, the covariance is 0 if $p = 1/2.$ For a fixed $n,\ Cov(\bar{X},\ S^2)$ has the maximum value $\sqrt{3}/(18n)$ when $6p = 3-\sqrt{3},$ or, $p \approx 0.21$ and the minimum value $-\sqrt{3}/(18n)$ when $6p = 3+\sqrt{3},$ or, $p \approx 0.79.$ Also $\lim_{n\to\infty} Cov(\bar{X},\ S^2) = 0.$

For instructional ease, the joint distribution of sample mean and variance for $n = 3$ is given below:

$(\bar{x}, s^2)$	$(0, 0)$	$(1/3, 1/3)$	$(2/3, 1/3)$	$(1, 0)$
$f(\bar{x}, s^2)$	q^3	$3pq^2$	$3p^2q$	p^3

One can easily check that $\bar{X}$ and S^2 are uncorrelated if $p = 1/2.$

In general, we remark that if $X_1,\ X_2,\ \cdots,\ X_n$ are independently and identically distributed, then $Cov(\bar{X}, S^2) = 0,$ iff the distribution is symmetric around mean μ and $\mu_3 = 0$ where $\mu_3 = E(X-\mu)^3.$

Example 3.2 Let $X_1,\ X_2,\ \cdots,\ X_n$ be independently and identically distributed observations of size $n \geq 2$ from continuous uniform distribution $U(\alpha, \beta)$. Then the sample mean $\bar{X}$ and the variance S^2 are uncorrelated.

Example 3.3 Let $X_1,\ X_2,\ \cdots,\ X_n$ be independently and identically distributed observations of size $n \geq 2$ from a discrete uniform distribution on the support $\{1,\ 2, \cdots, N\}$. Then the sample mean $\bar{X}$ and the variance S^2 are uncorrelated.

The following example from Mukhopadhyay and Son (2011) is related to Corollary 3.2 of this paper. They solved it without explicitly mentioning Corollary 3.2.

Example 3.4 Let the observations $X_1, X_2, \cdots, X_n$ be identically distributed with $\underset{\sim}{X} \sim N_n(\underset{\sim}{0}, \Sigma)$ where $\underset{\sim}{X} = (X_1, X_2, \cdots, X_n)',\ \Sigma = (\sigma_{ij})_{1\leq i\leq n, 1\leq j\leq n},\ \sigma_{ii} = \sigma^2$ and $\sigma_{ij} = \rho\sigma^2,\ (i \neq j)$ and $j = 1, 2, \cdots, n,$ where $-(n-1)^{-1} < \rho < 1,$ and $n = 4, 5, \cdots.$ The observations are not independently distributed unless $\rho = 0.$

Obviously, $\mu = \mu_3' = 0.$ Then $\mu_{21}' = 0$ (see Kotz, Balakrishnan and Johnson, 2000, p. 261) and $E(X_1 X_2 X_3) = 0$ (see Anderson, 1984, p. 49), so that $\gamma_{123} = 2[E(X_1 X_2^2) - E(X_1 X_2 X_3) - \mu\sigma^2]$ vanishes. Then by (3.14), $Cov(\bar{X}, S^2) = 0.$

4. Covariance Between Linear and Quadratic Functions

Let $\underset{\sim}{X}' = (X_1, X_2, \cdots X_n)$ be a vector of n jointly distributed random variables with respective means μ_j and variance σ_j^2 $(j = 1, 2, \cdots, n)$. Also let $L = \underset{\sim}{b}'\underset{\sim}{X} = \sum_{j=1}^{j=n} b_j X_j$ and $Q = \underset{\sim}{X}'A\underset{\sim}{X} = \sum_{i=1}^{i=n}\sum_{j=1}^{j=n} a_{ij} X_i X_j$ be respectively a linear and a quadratic form of X_j where $A = [a_{ij}]$ is an $n \times n$ symmetric matrix. Our objective is to determine a simple for the covariance $Cov(L, Q)$, using an orthogonal diagonalization of the matrix A. Indeed, since A is a symmetric matrix, there exists an orthogonal matrix C (that is, $C'C$ is the identity matrix) and a diagonal matrix $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ such that $C'AC = D$ where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A.

Theorem 4.1 Let $L = \underset{\sim}{b}'\underset{\sim}{X} = \sum_{j=1}^{j=n} b_j X_j$ and $Q = \underset{\sim}{X}'A\underset{\sim}{X} = \sum_{i=1}^{i=n}\sum_{j=1}^{j=n} a_{ij} X_i X_j$. Then

$$Cov(L, Q) = \sum_{i=1}^{i=n}\sum_{j=1}^{j=n} \underset{\sim}{b}'\underset{\sim}{c}_{.j}\lambda_j Cov(Y_i, Y_j^2), \tag{4.1}$$

where $\underset{\sim}{c}_{.j}$ is the j-th column of C and $\underset{\sim}{Y} = C'\underset{\sim}{X}$.

Proof. Define a vector $\underset{\sim}{X}' = (X_1, X_2, \cdots X_n)$ by $\underset{\sim}{Y} = C'\underset{\sim}{X}$. Then $L = \underset{\sim}{b}'C\underset{\sim}{Y} = \sum_{j=1}^{j=n} \underset{\sim}{b}'\underset{\sim}{c}_{.j} Y_j$ and $Q = (C\underset{\sim}{Y})'A(C\underset{\sim}{Y}) = \underset{\sim}{Y}'C'AC\underset{\sim}{Y} = \underset{\sim}{Y}'D\underset{\sim}{Y} = \sum_{j=1}^{j=n} \lambda_j Y_j^2$ where $\underset{\sim}{c}_{.j}$ is the j-th column of C. Then

$$Cov(L, Q) = E\left[\left(\sum_{j=1}^{j=n} \underset{\sim}{b}'\underset{\sim}{c}_{.j} Y_j\right)\left(\sum_{j=1}^{j=n} \lambda_j Y_j^2\right)\right] - E\left(\sum_{j=1}^{j=n} \underset{\sim}{b}'\underset{\sim}{c}_{.j} Y_j\right) E\left(\sum_{j=1}^{j=n} \lambda_j Y_j^2\right),$$

which can be written as $Cov(L, Q) = \sum_{i=1}^{i=n}\sum_{j=1}^{j=n} \underset{\sim}{b}'\underset{\sim}{c}_{.i}\lambda_j \left[E\left(Y_i Y_j^2\right) - E(Y_i)E\left(Y_j^2\right)\right]$, which is equivalent to (4.1).

Theorem 4.2 Let $Y = C'X$. Then

$$Cov(\bar{X}, S^2) = \frac{1}{(n-1)\sqrt{n}} \sum_{j=1}^{j=n-1} Cov(Y_n, Y_j^2). \tag{4.2}$$

Proof. Let $\underset{\sim}{b}' = (1, 1, \cdots, 1)$, $A = [a_{ij}]$ where $a_{ii} = 1 - (1/n)$, $i = 1, 2, \cdots, n$, $a_{ij} = -(1/n)$, $i = 1, 2, \cdots, n$, $j (\neq i) = 1, 2, \cdots, n$. Note that the matrix A can be written as $A = I - (1/n)J$ where J is a $n \times n$ matrix with all entries equal to 1. Also let $\underset{\sim}{c}_{.j}$ be the j-th column of

$$C = \begin{bmatrix} -1/\sqrt{1(2)} & -1/\sqrt{2(3)} & \cdots & -1/\sqrt{n(n-1)} & 1/\sqrt{n} \\ 1/\sqrt{1(2)} & -1/\sqrt{2(3)} & \cdots & -1/\sqrt{n(n-1)} & 1/\sqrt{n} \\ 0 & 2/\sqrt{2(3)} & \cdots & -1/\sqrt{n(n-1)} & 1/\sqrt{n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & (n-1)/\sqrt{n(n-1)} & 1/\sqrt{n} \end{bmatrix}.$$

It is easy to check that $L = \sum_{j=1}^{j=n} X_j$, $Q = \sum_{j=1}^{j=n} (X_j - \bar{X})^2$ and $C'C = I_n$. Then for $j = 1, 2, \cdots, n-1$, we have $A\underset{\sim}{c}_{.j} = (I - (1/n)J)\underset{\sim}{c}_{.j} = \underset{\sim}{c}_{.j}$, and for $j = n$, we have $A\underset{\sim}{c}_{.n} = (I - (1/n)J)\underset{\sim}{c}_{.n} = \underset{\sim}{0}$.

The equation $A\underset{\sim}{c}_{.j} = 1 \times \underset{\sim}{c}_{.j}$, $j = 1, 2, \cdots, n-1$, means that $\underset{\sim}{c}_{.j}$ is the eigenvector of the matrix A corresponding to the eigen value 1. Also the equation $A\underset{\sim}{c}_{.n} = 0 \times \underset{\sim}{c}_{.n}$, means that $\underset{\sim}{c}_{.n}$ is the eigenvector of the matrix A corresponding to the eigen value 0. We then have

$$C'AC = \begin{bmatrix} I_{n\times n} & \underset{\sim}{0}_{n-1} \\ \underset{\sim}{0}'_{n-1} & 0 \end{bmatrix}.$$

Since $Y = C'X$, and $C'C = I$, we have $\underset{\sim}{Y}'\underset{\sim}{Y} = (C\underset{\sim}{X})'C\underset{\sim}{X} = \underset{\sim}{X}'C'C\underset{\sim}{X} = \underset{\sim}{X}'\underset{\sim}{X}$. Also we have $Y_n = \sqrt{n}\bar{X}$ so that $(n-1)S^2 = \sum_{i=1}^{i=n} X_i^2 - n\bar{X}^2 = \sum_{i=1}^{i=n} Y_i^2 - Y_n^2 = \sum_{i=1}^{i=n-1} Y_i^2$. Then $(n-1)\sqrt{n}Cov(\bar{X},\ S^2) = Cov\left(Y_n,\ \sum_{i=1}^{i=n-1} Y_i^2\right)$. The proof is thus complete.

Acknowledgements The authors gratefully acknowledges the excellent research facilities provided by King Fahd University of Petroleum and Minerals, Saudi Arabia through the Project IN13105.

References

Fisher, R.A. (1925). Applications of Student's Distribution. *Metron*, 5, 90-104.

Geary, R.C. (1936). The distribution of the Student's ratio for the non-normal samples. *Supplement to the Journal of the Royal Statistical Society*, 3(2), 178-184.

Kang, K. and Goldsman, D. (1985). The correlation between mean and variance estimators. Proceedings of the 19985 Winter Simulation Conference, pp. 211-216. Edited by D. Gantz, G. Blais and S. Solomon.

Kang, K. and Goldsman, D. (1990). The correlation between mean and variance estimators in computer simulation. *IEE Transactions*, 22(1), 15-23.

Mukhopadhyay, N. (2005). Dependence and independence of the sample mean and variance in non-iid or non-normal cases and the role of some tests of independence. In *Recent Advances in Applied Probability*, IWAP 2002 Caracus Proceedings edited by Baeza-Yates, R.; Glaz, J. ; Husler, J. and Palacious, J.L. New York: Springer Verlag, pp.397-426.

Mukhopadhyay, N. and Son, M.S. (2011). On the covariance between the sample mean and variance. *Communications in Statistics – Theory and Methods*, 40, 1142-1148.

Student (1908). The probable error of a mean. *Biometrika*, 6(1), 1-25.

Zhang, L. (2007). Sample mean and variance: Their Covariance and their (In)depenedence. *American Statistician*, 61(2), 159-160.

File: RES-MATH-STAT \ COV-MEAN-VAR \ Cov-Mean-Var-S