ON THE STOCHASTIC INDEPENDENCE OF TWO SECOND-DEGREE POLYNOMIAL STATISTICS IN NORMALLY DISTRIBUTED VARIATES

By R. G. Laha

Indian Statistical Institute

A remarkable property of the normal law as proved by Craig [1] is that if $x_1, x_2, \cdots, x_n$ are n identically and independently distributed normal variates each with zero mean and unit variance, then the necessary and sufficient condition for the stochastic independence of two real homogeneous quadratic statistics $Q_1 = xAx'$ and $Q_2 = xBx'$ is that the matrix product $AB = 0$. The same theorem has also been proved independently by Hotelling [2], Sakamoto [5], Matusita [3], and Ogawa [4].

In the present paper we shall establish a corresponding theorem for the case of two second-degree polynomial statistics in normally distributed variates, and give some related results.

Theorem 1. *Let $x_1, x_2, \cdots, x_n$ be n independently and identically distributed normal variates each with zero mean and unit variance; then the necessary and sufficient condition that two real polynomial statistics of the second degree denoted by $P_1 = xAx' + lx'$ and $P_2 = xBx' + mx'$ are stochastically independent is that*

$$\text{(i) } AB = 0, \qquad \text{(ii) } lB = 0, \qquad \text{(iii) } mA = 0, \qquad \text{(iv) } lm' = 0.$$

Here, x, l, and m, respectively, represent the row-vectors $(x_1, x_2, \cdots, x_n)$, $(l_1, l_2, \cdots, l_n)$, and $(m_1, m_2, \cdots, m_n)$ and x', l', and m', as usual, represent their corresponding transposes and $A = (a_{ij})$ and $B = (b_{ij})$ are both real symmetric matrices of order n.

Proof of Sufficiency. Without any loss of generality we can write t_1 and t_2 in place of it_1 and it_2, respectively, so that the characteristic function of the joint distribution of P_1 and P_2 is given by

$$\phi(t_1, t_2) = E[\exp(t_1P_1 + t_2P_2)]. \tag{1}$$

Hence,

$$\begin{aligned}\phi(\tfrac{1}{2}t_1, \tfrac{1}{2}t_2) &= E[\exp(\tfrac{1}{2}t_1P_1 + \tfrac{1}{2}t_2P_2)]\\ &= |I - t_1A - t_2B|^{1/2}\\ &\quad \times \exp\{\tfrac{1}{8}(t_1l + t_2m)(I - t_1A - t_2B)^{-1}(t_1l + t_2m)'\}.\end{aligned} \tag{2}$$

Now, putting $t_2 = 0$ and $t_1 = 0$ alternatively in (2), we get

$$\phi(\tfrac{1}{2}t_1, 0) = |I - t_1A|^{1/2} \exp[\tfrac{1}{8}t_1^2 l(I - t_1A)^{-1}l'], \tag{3a}$$

$$\phi(0, \tfrac{1}{2}t_2) = |I - t_2B|^{1/2} \exp[\tfrac{1}{8}t_2^2 m(I - t_2B)^{-1}m'], \tag{3b}$$

Received May 27, 1955.

where $\phi(t_1, 0)$ and $\phi(0, t_2)$ represent the characteristic functions of the marginal distributions of P_1 and P_2, respectively.

When $AB = 0$, we get, after a little simplification,

$$|I - t_1A||I - t_2B| = |I - t_1A - t_2B|, \tag{4a}$$

$$(I - t_1A)^{-1} + (I - t_2B)^{-1} - I = (I - t_1A - t_2B)^{-1}, \tag{4b}$$

and

$$\begin{aligned}(t_1l + t_2m)(I - t_1A - t_2B)^{-1}(t_1l + t_2m)' - t_1^2l(I - t_1A)^{-1}l' - t_2^2m(I - t_2B)^{-1}m' \\ = t_1^2t_2lB(I - t_2B)^{-1}l' + t_1t_2^2m(I - t_1A)^{-1}Am' \\ + 2t_1^2t_2l(I - t_1A)^{-1}Am' + 2t_1t_2^2lB(I - t_2B)^{-1}m' \\ + 2t_1t_2lm'.\end{aligned} \tag{4c}$$

Thus, when $lB = 0$, $mA = 0$, $lm' = 0$, in addition to the condition $AB = 0$, the expression on the right-hand side of (4c) vanishes, yielding the relation

$$\begin{aligned}(t_1l + t_2m)(I - t_1A - t_2B)^{-1}(t_1l + t_2m)' \\ = t_1^2l(I - t_1A)^{-1}l' + t_2^2m(I - t_2B)^{-1}m'.\end{aligned} \tag{5}$$

Then using (4a) and (5) together in (2), (3a) and (3b), we get $\phi(t_1, t_2) = \phi(t_1, 0)\phi(0, t_2)$, establishing the stochastic independence of P_1 and P_2.

PROOF OF NECESSITY. Here it is given that the relation $\phi(t_1, t_2) = \phi(t_1, 0)\phi(0, t_2)$ holds identically for all real t_1 and t_2, so that from (2), (3a), and (3b), we have the relation

$$\begin{aligned}\exp\{\tfrac{1}{4}[(t_1l + t_2m)(I - t_1A - t_2B)^{-1}(t_1l + t_2m)' \\ -t_1^2l(I - t_1A)^{-1}l' - t_2^2m(I - t_2B)^{-1}m']\} = \frac{|I - t_1A||I - t_2B|}{|I - t_1A - t_2B|}.\end{aligned} \tag{6}$$

Thus, from (6) we see that the relation

$$\exp[P(it_1, it_2)/Q(it_1, it_2)] = R(it_1, it_2)/S(it_1, it_2)$$

holds identically for all real t_1 and t_2, where P, Q, R, and S are polynomials in t_1 and t_2.

But it can be easily proved that in such a case the rational functions P/Q and R/S are constants. Hence, (6) gives two conditions:

$$|I - t_1A||I - t_2B| = C_1\cdot|I - t_1A - t_2B|, \tag{7a}$$

$$\begin{aligned}(t_1l + t_2m)(I - t_1A - t_2B)^{-1}(t_1l + t_2m)' \\ - t_1^2l(I - t_1A)^{-1}l' - t_2^2m(I - t_2B)^{-1}m' = C_2\end{aligned} \tag{7b}$$

to be satisfied for all t_1 and t_2, where C_1 and C_2 are constants.

But, putting $t_1 = t_2 = 0$ in (7a) and (7b), it follows that $C_1 = 1$ and $C_2 = 0$, so that we have

$$|I - t_1A| \, |I - t_2B| = |I - t_1A - t_2B|, \tag{8a}$$

$$\begin{aligned}(t_1l + t_2m)(I - t_1A - t_2B)^{-1}(t_1l + t_2m)' \\ - t_1^2l(I - t_1A)^{-1}l' - t_2^2m(I - t_2B)^{-1}m' = 0.\end{aligned} \tag{8b}$$

It has been already proved in [1], [2], [3], [4], and [5] that if (8a) holds identically for all t_1 and t_2, then $AB = 0$.

Now, by virtue of the condition $AB = 0$, the left-hand side of (8b) simplifies to the expression (4c). Hence, we have

$$\begin{aligned}t_1^2t_2lB(I - t_2B)^{-1}l' + t_1t_2^2m(I - t_1A)^{-1}Am' \\ + 2t_1^2t_2l(I - t_1A)^{-1}Am' + 2t_1t_2^2lB(I - t_2B)^{-1}m' \\ + 2t_1t_2lm' = 0,\end{aligned} \tag{8c}$$

holding identically for all t_1 and t_2. Then restricting the values of t_1 and t_2 to the neighbourhoods of the origin $|t_1| < 1/\alpha$ and $|t_2| < 1/\beta$, where α and β denote the largest of the absolute values of the latent roots of the matrices A and B, respectively, we have the power series expansion

$$(I - t_1A)^{-1} = I + t_1A + t_1^2A^2 + \cdots, \tag{9a}$$

$$(I - t_2B)^{-1} = I + t_2B + t_2^2B^2 + \cdots. \tag{9b}$$

Substituting the expressions on the right-hand sides of (9a) and (9b) in (8c), above, and collecting the coefficients of t_1t_2 and $t_1^2t_2^2$, we get

$$lm' = 0, \tag{10a}$$

$$lBBl' + mAAm' = 0. \tag{10b}$$

The elements of A, B, l, and m being all real, (10b) at once gives $lB = 0$ and $mA = 0$. Again, when $lB = 0$, $mA = 0$, and $lm' = 0$, (8c), above, is satisfied for all t_1 and t_2, which completes the proof.

The extension to the correlated normal variates is also simple. Let $x = (x_1, x_2, \cdots, x_n)$ be n-variate normal with mean vector zero and the variance-covariance matrix Σ. Then the necessary and sufficient condition that $P_1 = xAx' + lx'$ and $P_2 = xBx' + mx'$ are stochastically independent is that

(i) $A\Sigma B = 0$, (ii) $l\Sigma B = 0$, (iii) $m\Sigma A = 0$, (iv) $l\Sigma m' = 0$.

The proof follows by using a real nonsingular linear transformation $y = xT$ such that $TT' = \Sigma^{-1}$. Hence, using the above transformation, we have $P_1 = yA_0y' + l_0y'$ and $P_2 = yB_0y' + m_0y'$, where $A = TA_0T'$, $B = TB_0T'$, $l = l_0T'$, and $m = m_0T'$, and further $y_1, y_2, \cdots, y_n$ are independently normally distributed, each with zero mean and unit variance. Then using the above theorem, we get the set of necessary and sufficient conditions as

(11) (i) $A_0B_0 = 0$, (ii) $l_0B_0 = 0$, (iii) $m_0A_0 = 0$, and (iv) $l_0m_0' = 0$.

Next, rewriting the conditions in (11) in terms of A, B, l, and m, and using the relation $TT' = \Sigma^{-1}$, we get

$$\text{(i) } A\Sigma B = 0, \qquad \text{(ii) } l\Sigma B = 0, \qquad \text{(iii) } m\Sigma A = 0, \qquad \text{(iv) } l\Sigma m' = 0.$$

COROLLARY I. (Extension to the noncentral case). *Let $x_1, x_2, \cdots, x_n$ be n independent normal variates distributed with means $\mu_1, \mu_2, \cdots, \mu_n$, but having the same variance, say unity. Then the necessary and sufficient condition for the stochastic independence of two second degree polynomial statistics $P_1 = xAx' + lx'$ and $P_2 = xBx' + mx'$ is that*

$$\text{(i) } AB = 0, \qquad \text{(ii) } lB = 0, \qquad \text{(iii) } mA = 0, \qquad \text{(iv) } lm' = 0.$$

PROOF. Let us take $y = x - \mu$, where

$$y = (y_1, y_2, \cdots, y_n), \qquad x = (x_1, x_2, \cdots, x_n), \mu = (\mu_1, \mu_2, \cdots, \mu_n).$$

Then we have

$$P_1 = xAx' + lx' = yAy' + (2\mu A + l)y' + \mu A\mu' + l\mu',$$

$$P_2 = xBx' + mx' = yBy' + (2\mu B + m)y' + \mu B\mu' + m\mu'.$$

Now, $y_1, y_2, \cdots, y_n$ are also distributed independently normally, each having zero mean and unit variance, and the proof follows from the above theorem, simply replacing l by $2\mu A + l$ and m by $2\mu B + m$. The corresponding extension to the correlated case when the variates are distributed with an artibrary mean vector $(\mu_1, \mu_2, \cdots, \mu_n)$ and variance covariance matrix Σ also follows immediately.

COROLLARY II.[1] *Let $x_1, x_2, \cdots, x_n$ be n identically and independently distributed normal variates each having zero mean and unit variance. If two real polynomial statistics of the second degree denoted by $P_1 = xAx' + lx'$ and $P_2 = xBx' + mx'$ are stochastically independent, then there always exists an orthogonal transformation given by $y = xP$, reducing simultaneously both $P_1(x_1, x_2, \cdots, x_n)$ to $P_1'(y_1, y_2, \cdots, y_k)$ and $P_2(x_1, x_2, \cdots, x_n)$ to $P_2'(y_{k+1}, y_{k+2}, \cdots, y_n)$ such that P_1' and P_2' do not contain any common variate.*

In this connection it is interesting to note that a more general and difficult problem has been suggested by Prof. Yu. V. Linnik during his recent seminar at the Indian Statistical Institute in Calcutta. It is his conjecture that if two polynomial statistics $P_1(x_1, x_2, \cdots, x_n)$ and $P_2(x_1, x_2, \cdots, x_n)$ in identically and independently distributed normal variates $x_1, x_2, \cdots, x_n$ are stochastically independent, then there exists an orthogonal transformation $y = xP$, reducing simultaneously both $P_1(x_1, x_2, \cdots, x_n)$ to $P_1'(y_1, y_2, \cdots, y_k)$ and $P_2(x_1, x_2, \cdots, x_n)$ to $P_2'(y_{k+1}, y_{k+2}, \cdots, y_n)$ such that the new polynomials P_1' and P_2' do not contain any common variate; that is, we can "unlink" the poly-

---

[1] While this paper was in press, the author has learned in a communication from Prof. Yu. V. Linnik that the results contained in Corollary II are also obtained independently by Prof. A. A. Zinger of the University of Leningrad. But his method of proof is not known to the author.

nomials in such a case. Here we give a partial solution to this problem when both the polynomials are of the second degree. We note further that the above corollary generalizes a corresponding result due to Hotelling [2] on the stochastic independence of two real homogeneous quadratic statistics.

PROOF. From the above theorem, it follows that when P_1 and P_2 are stochastically independent, we have

$$AB = 0, \qquad lB = 0, \qquad mA = 0, \quad \text{and} \quad lm' = 0.$$

Let $\alpha_1, \alpha_2, \cdots, \alpha_r$ and $\beta_1, \beta_2, \cdots, \beta_s$ denote the non-zero latent roots of the matrices A and B, respectively, such that $r + s \leqq n$. Now there exists an orthogonal matrix C such that $CAC' = D$ and $CBC' = E$, where

$$D = \begin{array}{cc} \begin{array}{cc} r & n-r \end{array} & \\ \begin{pmatrix} D_\alpha & 0 \\ 0 & 0 \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \end{array}, \qquad E = \begin{array}{cc} \begin{array}{cc} r & n-r \end{array} & \\ \begin{pmatrix} E_1 & E_2 \\ E_2' & F \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \end{array},$$

D_α being the diagonal matrix consisting of the non-zero latent roots $\alpha_1, \alpha_2, \cdots, \alpha_r$ given by

$$D_\alpha = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_r \end{pmatrix}.$$

The matrix C being orthogonal, the relation (8a) reduces to

$$|I - t_1D|\,|I - t_2E| = |I - t_1D - t_2E|. \tag{12}$$

Next, equating the coefficients of t_1^r on both sides of (12), we get

$$|I - t_2E| = |I - t_2F|, \tag{13}$$

holding for all t_2, where F is the symmetric matrix of order $n - r$ as given above. Now, from (13), it can be shown easily that the non-zero latent roots of the matrix F are also given by $\beta_1, \beta_2, \cdots, \beta_s$. Hence, there exists an orthogonal matrix C_0 of order $n - r$ such that

$$C_0FC_0' = \begin{array}{cc} \begin{array}{cc} s & n-r-s \end{array} & \\ \begin{pmatrix} E_\beta & 0 \\ 0 & 0 \end{pmatrix} & \begin{array}{l} s \\ n-r-s \end{array} \end{array},$$

where E_β is the diagonal matrix of order s formed by the non-zero latent roots $\beta_1, \beta_2, \cdots, \beta_s$ given by

$$E_\beta = \begin{pmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_s \end{pmatrix}.$$

Let us now put

$$C_1 = \begin{array}{cc} \begin{matrix} r & n-r \end{matrix} & \\ \begin{pmatrix} I & 0 \\ 0 & C_0 \end{pmatrix} & \begin{matrix} r \\ n-r \end{matrix} \end{array}$$

and write $C_2 = C_1C$. Then it follows that C_2 is an orthogonal matrix and further that

$$C_2AC_2' = \begin{array}{cc} \begin{matrix} r & n-r \end{matrix} & \\ \begin{pmatrix} D_\alpha & 0 \\ 0 & 0 \end{pmatrix} & \begin{matrix} r \\ n-r \end{matrix} \end{array}$$

and

$$C_2BC_2' = \begin{array}{cc} \begin{matrix} r & \qquad n-r \end{matrix} & \\ \begin{pmatrix} E_1 & E_2C_0' \\ C_0E_2' & \begin{array}{cc} \begin{matrix} s & n-r-s \end{matrix} & \\ \begin{pmatrix} E_\beta & 0 \\ 0 & 0 \end{pmatrix} & \begin{matrix} s \\ n-r-s \end{matrix} \end{array} \end{pmatrix} & \begin{matrix} r \\ n-r \end{matrix} \end{array}.$$

Let us denote C_2BC_2' by G, then it follows that the latent roots of G are the same as those of B. But since the matrix G is symmetric, the sum of the squares of its elements is equal to the sum of the squares of its non-zero latent roots and hence equal to $\sum_{j=1}^{s}\beta_j^2$. Thus it follows that $E_1 = 0$ and $E_2 = 0$. Hence, we have proved the existence of an orthogonal matrix C_2, reducing simultaneously both the matrices A and B to their canonical forms such that

$$C_2AC_2' = D_0 = \begin{array}{cc} \begin{matrix} r & s & n-r-s \end{matrix} & \\ \begin{pmatrix} D_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{matrix} r \\ s \\ n-r-s \end{matrix} \end{array}$$

and

$$C_2BC_2' = E_0 = \begin{array}{cc} \begin{matrix} r & s & n-r-s \end{matrix} & \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & E_\beta & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{matrix} r \\ s \\ n-r-s \end{matrix} \end{array},$$

where D_α and E_β are defined above.

Let us now suppose that this orthogonal transformation, defined by $x = yC_2$, reduces the vectors l and m to λ and μ, respectively, such that $\lambda = lC_2'$ and $\mu = mC_2'$. Then the conditions $lB = 0$, $mA = 0$, and $lm' = 0$ give us $\lambda E_0 = 0$, $\mu D_0 = 0$, and $\lambda\mu' = 0$, respectively. Thus, from the form of D_0 and E_0, it follows that the vectors λ and μ should be of the forms

$$\lambda = (\overset{r}{\lambda_1, \lambda_2, \cdots, \lambda_r}; \overset{s}{0, 0, \cdots, 0}; \overset{n-r-s}{\lambda_{r+s+1}, \cdots, \lambda_n})$$

$$\mu = (\overset{r}{0, 0, \cdots, 0}; \overset{s}{\mu_1, \mu_2, \cdots, \mu_s}; \overset{n-r-s}{\mu_{r+s+1}, \cdots, \mu_n}),$$

where the elements satisfy the relation

$$\sum_{j=r+s+1}^{n} \lambda_j\mu_j = 0.$$

At first we note that if $r + s = n$, the above orthogonal transformation defined by $x = yC_2$ reduces simultaneously both

$$P_1 = xAx' + lx' \quad \text{to} \quad P_1' = \sum_{j=1}^{r} \alpha_j y_j^2 + \sum_{j=1}^{r} \lambda_j y_j$$

and

$$P_2 = xBx' + mx' \quad \text{to} \quad P_2' = \sum_{j=1}^{n-r} \beta_j y_{r+j}^2 + \sum_{j=1}^{n-r} \mu_j y_{r+j}$$

such that the new polynomials P_1' and P_2' have no common variate, which completes the proof.

But if $r + s < n$, we take

$$C_3 = \begin{array}{cccc} r & s & n-r-s & \\ \left(\begin{array}{c} I \\ 0 \\ 0 \end{array}\right. & \begin{array}{c} 0 \\ I \\ 0 \end{array} & \left.\begin{array}{c} 0 \\ 0 \\ C_4 \end{array}\right) & \begin{array}{l} r \\ s \\ n-r-s \end{array} \end{array}$$

where C_4 is an orthogonal matrix of order $n - r - s$ such that its first two row vectors are given by

$$(\lambda_{r+s+1}/\lambda_0, \cdots, \lambda_n/\lambda_0) \quad \text{and} \quad (\mu_{r+s+1}/\mu_0, \cdots, \mu_n/\mu_0),$$

where

$$\lambda_0^2 = \sum_{j=r+s+1}^{n} \lambda_j^2 \quad \text{and} \quad \mu_0^2 = \sum_{j=r+s+1}^{n} \mu_j^2 .$$

Thus, if we define the above transformation by $x = zP$, where $P = C_3C_2$ is an orthogonal matrix, it follows easily that this transformation reduces simultaneously both

$$P_1 = xAx' + lx' \quad \text{to} \quad P_1' = \sum_{j=1}^{r} \alpha_j z_j^2 + \sum_{j=1}^{r} \lambda_j z_j + \lambda_0 z_{r+s+1}$$

and

$$P_2 = xBx' + mx' \quad \text{to} \quad P_2' = \sum_{j=1}^{s} \beta_j z_{r+j}^2 + \sum_{j=1}^{s} \mu_j z_{r+j} + \mu_0 z_{r+s+2}$$

such that P_1' and P_2' do not have any common variate. Hence the proof.

The proof of the above corollary for the non-central normal case is also immediate.

REFERENCES

[1] A. T. Craig, "Note on the independence of certain quadratic forms," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 195–197.

[2] H. Hotelling, "Note on a matric theorem of A. T. Craig," *Ann. Math. Stat.*, Vol. 15 (1944), pp. 427–429.

[3] K. Matusita, "Note on the independence of certain statistics," *Ann. Ins. Stat. Math.*, Tokyo, Vol. 1 (1949), pp. 79–82.

[4] J. Ogawa, "On the independence of bilinear and quadratic forms of a random sample from a normal population," *Ann. Ins. Stat. Math.*, Tokyo, Vol. 1 (1949), pp. 83–108.

[5] H. Sakamoto, "On the independence of two statistics," *Res. Mem. Inst. Stat. Math.*, Tokyo, Vol. 1 (1944), pp. 1–25.