

*On the Independence of Quadratic Forms in a Non-Central Normal System**

By Junjiro OGAWA

The problem on the independence of quadratic forms in a normal system originates from the famous treatise of W. G. Cochran¹⁾. Cochran proved the following theorem: "Let x_1, x_2, \dots, x_n be independently distributed according to the identical normal law $N(0, 1)$ ²⁾, q_1, q_2, \dots, q_m being m quadratic forms of x_1, x_2, \dots, x_n , and their ranks r_1, r_2, \dots, r_m respectively. If $\sum_1^m q_j = \sum_1^n x_i^2$, then the necessary and sufficient condition for q_1, q_2, \dots, q_m to be independent statistically is that $\sum_1^m r_j = n$. When these conditions are satisfied, then q_1, q_2, \dots, q_m are distributed independently according to the chi-square distributions of degrees of freedom r_1, r_2, \dots, r_m respectively". Then, in 1940, W. G. Madow³⁾ proved the generalization of this theorem for the non-central case, and he obtained the same condition.

On the other hand, A. T. Craig⁴⁾, H. Hotelling⁵⁾ and H. Sakamoto⁶⁾ have extended the theorem in the other direction, their point being as follows: Let x_1, \dots, x_n be independently distributed according to the normal law $N(0, 1)$ and written as a vector $\xi = (x_1, \dots, x_n)$, and furthermore let A and B be two real symmetric matrices, then the necessary and sufficient condition for two quadratic forms $q_1 = \xi A \xi'$ and $q_2 = \xi B \xi'$ to be independent statistically is $AB = 0$. But their proofs were insufficient, and K. Matsushita⁷⁾ and we⁸⁾ gave the complete proofs.

In this note we shall generalize the last theorem for the non-central case, and prove the following two theorems and show one example of their applications.

§ 1. THEOREMS.

Theorem I. Let x_1, x_2, \dots, x_n be normally and independently distributed with means a_1, a_2, \dots, a_n respectively and with the common variance unity. If we denote n random variables x_1, x_2, \dots, x_n by a vector notation $\xi = (x_1, x_2, \dots, x_n)$, and make two quadratic forms $q_1 =$

$\xi A \xi'$ and $q_2 = \xi B \xi'$, where A and B are real symmetric matrices, then the necessary and sufficient condition of their statistical independence is

$$AB = 0. \quad (1)$$

PROOF⁹⁾: In the sequel we shall write α for the n -dimensional vector $(\alpha_1, \alpha_2, \dots, \alpha_n)$, which is neither proportional to unit vector nor to zero vector. The moment-generating function $\varphi(t_1, t_2)$ of the joint distribution of q_1 and q_2 is given as follows:

$$\varphi(t_1, t_2) = (2\pi)^{-\frac{n}{2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[t_1 q_1 + t_2 q_2 - \frac{1}{2} \sum_1^n (x_i - \alpha_i)^2 \right] dx_1 dx_2 \dots dx_n.$$

In vector notations,

$$\begin{aligned} \varphi(t_1, t_2) &= (2\pi)^{-\frac{n}{2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \xi (E - 2t_1 A - 2t_2 B) \xi' - 2\alpha \xi' + \alpha \alpha' \right] \\ &\quad dx_1 dx_2 \dots dx_n \\ &= (2\pi)^{-\frac{n}{2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} (\xi 1) \begin{pmatrix} E - 2t_1 A - 2t_2 B & -\alpha' \\ -\alpha & \alpha \alpha' \end{pmatrix} \begin{pmatrix} \xi' \\ 1 \end{pmatrix} \right] dx_1 dx_2 \dots dx_n, \end{aligned}$$

where E is the unit matrix of degree n and primes mean the transposed vectors.

We denote the matrix $E - 2t_1 A - 2t_2 B$ by C for a system of fixed values of t_1 and t_2 for which the matrix is positive-definite and operate the variates transformation represented by

$$(\xi 1) = (\eta 1) \begin{pmatrix} E & 0 \\ b & 1 \end{pmatrix},$$

where

$$\eta = (\eta_1, \eta_2, \dots, \eta_n)$$

and

$$b = (b_1, b_2, \dots, b_n) = \alpha C^{-1},$$

then

$$\begin{aligned} (\xi 1) \begin{pmatrix} C & -\alpha' \\ -\alpha & \alpha \alpha' \end{pmatrix} \begin{pmatrix} \xi' \\ 1 \end{pmatrix} &= (\eta 1) \begin{pmatrix} E & 0 \\ b & 1 \end{pmatrix} \begin{pmatrix} C & -\alpha' \\ -\alpha & \alpha \alpha' \end{pmatrix} \begin{pmatrix} E & b' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta' \\ 1 \end{pmatrix} \\ &= (\eta 1) \begin{pmatrix} C & 0 \\ 0 & -\alpha C^{-1} \alpha' + \alpha \alpha' \end{pmatrix} \begin{pmatrix} \eta' \\ 1 \end{pmatrix}, \end{aligned}$$

and the Jacobian of this transformation is unity.

Therefore, we have

$$\begin{aligned}\varphi(t_1, t_2) &= (2\pi)^{-\frac{n}{2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} (\mathbf{y} | \mathbf{1}) \begin{pmatrix} C & 0 \\ 0 & -\alpha C^{-1} \alpha' + \alpha \alpha' \end{pmatrix} \begin{pmatrix} \mathbf{y}' \\ \mathbf{1} \end{pmatrix} \right] dy_1 dy_2 \dots dy_n \\ &= |C|^{-\frac{1}{2}} \exp \left[\frac{1}{2} \alpha (C^{-1} - E) \alpha' \right].\end{aligned}$$

i. e.

$$\varphi(t_1, t_2) = |E - 2t_1A - 2t_2B|^{-\frac{1}{2}} \exp \left[\frac{1}{2} \alpha \{ (E - 2t_1A - 2t_2B)^{-1} - E \} \alpha' \right].$$

The necessary and sufficient condition of the statistical independence of q_1 and q_2 , as is well known, is given by

$$\varphi(t_1, t_2) = \varphi(t_1, 0) \varphi(0, t_2),$$

which becomes after some calculations,

$$\begin{aligned}& \frac{|E - 2t_1A| |E - 2t_2B|}{|E - 2t_1A - 2t_2B|} \\ &= \exp \left\{ \alpha \left[(E - 2t_1A)^{-1} + (E - 2t_2B)^{-1} - (E - 2t_1A - 2t_2B)^{-1} - E \right] \alpha' \right\} \quad (2)\end{aligned}$$

So we have only to show the equivalence of (1) and (2).

First, (1) implies (2): If $AB=0$, then A and B will be brought into diagonal forms simultaneously by means of the same orthogonal transformation, say P , and so, we shall have no loss of generality by assuming that

$$PAP' = \begin{pmatrix} \alpha_1 & & & 0 \\ & \ddots & & \\ & & \alpha_r & \\ & & & \dots \\ & & 0 & \\ 0 & & & \dots \\ & & & & 0 \end{pmatrix} \quad \text{and} \quad PBP' = \begin{pmatrix} 0 & & & 0 \\ & \ddots & & \\ & & 0 & \\ & & & \dots \\ & & \beta_1 & \\ & & \dots & \\ & & & \beta_s \\ & & 0 & \\ & & & \dots \\ 0 & & & & 0 \end{pmatrix} \quad (4)$$

where r and s are the respective ranks of A and B .

If we operate the orthogonal transformation P on both sides of the equation (2), then the left-hand side of the equation is clearly seen to be unity, therefore, to obtain the equality of both sides, it suffices to show that the matrix in the curled bracket on the right-hand side is zero matrix. It is certainly true, because

$$\begin{aligned}
 & (E-2t_1A)^{-1} + (E-2t_2B)^{-1} - (E-2t_1A-t_2B)^{-1} - E \\
 = & P' \left[(E-2t_1PAP')^{-1} + (E-2t_2PBP')^{-1} - (E-2t_1PAP' - 2t_2PBP')^{-1} - E \right] P \\
 = & P' \left(\begin{array}{c} \frac{1}{1-2t_1\alpha_1} \quad 0 \\ \vdots \\ \frac{1}{1-2t_1\alpha_r} \\ \vdots \\ 0 \quad \left. \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right\} n-r \end{array} \right) + \left(\begin{array}{c} \left. \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right\} r \quad 0 \\ \frac{1}{1-2t_2\beta_1} \\ \vdots \\ \frac{1}{1-2t_2\beta_s} \\ \vdots \\ 0 \quad \left. \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right\} n-r-s \end{array} \right) \\
 - & \left(\begin{array}{c} \frac{1}{1-2t_1\alpha_1} \quad 0 \\ \vdots \\ \frac{1}{1-2t_1\alpha_r} \\ \frac{1}{1-2t_2\beta_1} \\ \vdots \\ \frac{1}{1-2t_2\beta_s} \\ \vdots \\ 0 \quad \left. \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right\} n-r-s \end{array} \right) - \left(\begin{array}{c} 1 \\ \vdots \\ \left. \begin{array}{c} \vdots \\ \vdots \end{array} \right\} n \\ 1 \end{array} \right) P=0.
 \end{aligned}$$

Second, (2) implies (1): (2) is an equation of the form

$$F(t_1, t_2) = e^{G(t_1, t_2)}, \tag{3}$$

where

$$F(t_1, t_2) = \frac{|E-2t_1A| |E-2t_2B|}{|E-2t_1A-2t_2B|},$$

and

$$G(t_1, t_2) = \alpha \left\{ (E-2t_1A)^{-1} + (E-2t_2B)^{-1} - (E-2t_1A-2t_2B)^{-1} - E \right\} \alpha',$$

so they are rational functions of t_1 and t_2 which are analytic at the origin.

Let $F(t_1, t_2)$ and $G(t_1, t_2)$ be analytic in the neighbourhood of the origin, $|t_1| < \varepsilon$, $|t_2| < \varepsilon$. If we fix a value of t_2 such that $|t_2| < \varepsilon$, then $G(t_1, t_2)$ is a rational function of t_1 , which has no poles in the

whole complex t_1 -plane including ∞ . Otherwise, a pole of $G(t_1, t_2)$ would be an essential singularity of $e^{G(t_1, t_2)}$, which contradicts (3), where $F(t_1, t_2)$ was a rational function of t_1 . So $G(t_1, t_2)$ is a constant in t_1 , provided a value of t_2 is fixed such that $|t_2| < \varepsilon$. In the same way we can prove that $G(t_1, t_2)$ is a constant in t_2 , provided a value of t_1 is fixed such that $|t_1| < \varepsilon$. Hence $G(t_1, t_2)$ is a constant in the domain $|t_1| < \varepsilon, |t_2| < \varepsilon$. From the equation (3), the same is true for $F(t_1, t_2)$, so that

$$F(t_1, t_2) = F(0, 0) = 1$$

i. e.

$$|E - 2t_1A - 2t_2B| = |E - 2t_1A| |E - 2t_2B| \tag{4}$$

holds identically in the domain $|t_1| < \varepsilon, |t_2| < \varepsilon$. (4) is equivalent to (1)¹⁰⁾.

Theorem 2. Let $x_\nu = (x_{1\nu}, x_{2\nu}, \dots, x_{k\nu}) \nu = 1, 2, \dots, n$ be a random sample of size n drawn from a k -dimensional normal population

$$(2\pi)^{-\frac{k}{2}} |\Lambda|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \sum_{i,j} \lambda^{ij} (x_i - a_i)(x_j - a_j) \right],$$

where $\Lambda = (\lambda_{ij})$ is the moment-matrix of x_1, x_2, \dots, x_k and (λ^{ij}) is its inverse. If q_1 and q_2 are the quadratic forms in $x_{i\nu}, i = 1, 2, \dots, k, \nu = 1, 2, \dots, n$, then the criterion of the statistical independence of q_1 and q_2 is given by

$$A \cdot (\Lambda \times E) \cdot B = 0,$$

where A and B are the respective matrices of q_1 and q_2 and $\Lambda \times E$ denotes the Kronecker's product of Λ and the unit matrix of degree n .

PROOF: As in the previous paper¹¹⁾, if we make use of the non-singular variate transformation, which brings Λ into unit matrix, then the Theorem 2 reduces to Theorem 1.

§ 2. Application.

D. S. Villars¹²⁾ has recently treated the significance test and estimation of exponential regression. His point being as follows: Consider a variate z , whose distribution for a given value of a fixed variate, t , is:

$$f(z|t) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} (z - a + b e^{-kt})^2 \right\}, \tag{5}$$

where a, b and k are real-valued parameters. The regression of z on

t is exponential, for it follows from (5) that the expected value of z , given t , is:

$$E(z|t) = a - b e^{-kt} \quad (6)$$

On the basis of a random sample $O_N(z_1, t_1; z_2, t_2, \dots, z_N, t_N)$, it is required to test whether $k=0$ or ∞ .

When N is an even integer (≥ 6), and times t_1, t_2, \dots, t_N at which measurements of z are made are arranged such that

$$t_{2\alpha} - t_{2\alpha-1} = \Delta, \text{ a constant } \left(\alpha = 1, 2, \dots, n = \frac{N}{2} \right), \quad (7)$$

the odd time intervals $t_3 - t_2, t_5 - t_4, \dots$ do not have to be equal.

Let $x_\alpha = z_{2\alpha-1}$ and $y_\alpha = z_{2\alpha}$ for $\alpha = 1, 2, \dots, n$. From (5) and (7), it follows that n pairs x_α, y_α are normally and independently distributed with common variance σ^2 , and that

$$\nu_\alpha = h + m \mu_\alpha, \alpha = 1, 2, \dots, n, \quad (8)$$

where $\mu_\alpha = E(x_\alpha)$, $\nu_\alpha = E(y_\alpha)$, $h = a(1 - e^{-k\Delta})$, and $m = e^{-k\Delta}$.

And then, Villars proposed¹³⁾ the criterion

$$F' = \frac{S_{xx} + 2mS_{xy} + m^2S_{yy}}{m^2S_{xx} - 2mS_{xy} + S_{yy}} \quad (9)$$

for testing the null-hypothesis, where

$$S_{xx} = \sum_{\alpha=1}^n (x_\alpha - \bar{x})^2, \quad S_{xy} = \sum_{\alpha=1}^n (x_\alpha - \bar{x})(y_\alpha - \bar{y}), \quad S_{yy} = \sum_{\alpha=1}^n (y_\alpha - \bar{y})^2,$$

$$\bar{x} = \frac{1}{n} \sum_{\alpha=1}^n x_\alpha, \quad \bar{y} = \frac{1}{n} \sum_{\alpha=1}^n y_\alpha.$$

In this case,

$$q_1 = S_{xx} + 2mS_{xy} + m^2S_{yy} \text{ and } q_2 = m^2S_{xx} - 2mS_{xy} + S_{yy}$$

are both non-negative quadratic forms of a non-central normal system, and so their statistical independence can be judged by the Theorem 1 of the previous section.

The non-central normal system to be considered here is:

$$(2\pi)^{-n} \sigma^{-2n} \exp \left[-\frac{1}{2\sigma^2} \left\{ \sum_{\alpha=1}^n (x_\alpha - \mu_\alpha)^2 + \sum_{\alpha=1}^n (y_\alpha - \nu_\alpha)^2 \right\} \right] \prod dx_\alpha \prod dy_\alpha. \quad (10)$$

Here we make use of the so-called "Helmert's orthogonal transformation", represented by the matrix

$$P = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{n-1}{n} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} \\ 0 & \frac{\sqrt{n-2}}{\sqrt{n-1}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{1}{\sqrt{(n-1)(n-2)}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2,1}} \end{pmatrix}$$

and write

$$\begin{aligned} (\xi_1, \xi_2, \dots, \xi_n) &= (x_1, x_2, \dots, x_n)P', \\ (\eta_1, \eta_2, \dots, \eta_n) &= (y_1, y_2, \dots, y_n)P', \\ (\mu_1', \mu_2', \dots, \mu_n') &= (\mu_1, \mu_2, \dots, \mu_n)P', \end{aligned}$$

and

$$(\nu_1', \nu_2', \dots, \nu_n') = (\nu_1, \nu_2, \dots, \nu_n)P'.$$

The Jacobian of this transformation being unity, so (10) is transformed into

$$(2\pi)^{-n} \sigma^{-2n} \exp \left[-\frac{1}{2\sigma^2} \left\{ \sum_{\alpha=1}^n (\xi_\alpha - \mu_\alpha')^2 + \sum_{\alpha=1}^n (\eta_\alpha - \nu_\alpha')^2 \right\} \right] \prod d\xi_\alpha \prod d\eta_\alpha,$$

Further q_1 and q_2 are transformed into

$$q_1 = \sum_{\alpha=2}^n \xi_\alpha^2 + 2m \sum_{\alpha=2}^n \xi_\alpha \eta_\alpha + m^2 \sum_{\alpha=2}^n \eta_\alpha^2$$

and

$$q_2 = m^2 \sum_{\alpha=2}^n \xi_\alpha^2 - 2m \sum_{\alpha=2}^n \xi_\alpha \eta_\alpha + \sum_{\alpha=2}^n \eta_\alpha^2$$

respectively. So the matrices of q_1 and q_2 are transformed into

$$PAP' = \begin{pmatrix} 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \\ 0 & 1 \cdots 0 & 0 & m \cdots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots 1 & 0 & 0 \cdots m \\ 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \\ 0 & m \cdots 0 & 0 & m^2 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots m & 0 & 0 \cdots m^2 \end{pmatrix} \quad \text{and} \quad PBP' = \begin{pmatrix} 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \\ 0 & m^2 \cdots 0 & 0 & -m \cdots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots m^2 & 0 & 0 \cdots -m \\ 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \\ 0 & -m \cdots 0 & 0 & 1 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots -m & 0 & 0 \cdots 1 \end{pmatrix}$$

respectively, hence, for any real value of m , it is readily seen that

$$AB = 0.$$

Therefore, q_1 and q_2 are independent statistically. In addition, $q_2/\sigma^2(1+m^2)$ is distributed according to the χ^2 -distribution of degrees of freedom $n-1$, and $q_1/\sigma^2(1+m^2)$ is distributed according to the non-central χ^2 -distribution¹⁴⁾. Therefore, from (9), F' is distributed according to the so-called "non-central F' distribution".

(Received June 24, 1950)

References and Notes

1) W. G. Cochran, The Distribution of Quadratic Forms in a Normal System, with Applications to the Analysis of Covariance, Proc. Camb. Phil. Soc., Vol. 30, 1934.

2) $N(0, 1)$ denotes the normal distribution law with mean 0 and unit variance, i. e., the standard normal distribution.

3) W. G. Madow, The Distribution of Quadratic Forms in Non-central Normal Random Variable, Ann. Math. Statist. Vol. II, 1940.

W. G. Madow, On a Source of Downwards Bias in the Analysis of Variance and Covariance, Ann. Math. Statist., Vol. 19 (1948) pp. 351-359.

4) A. T. Craig, On the Independence of Certain Estimates of Variance, Ann. Math. Statist., Vol. 9, 1938.

A. T. Craig, Note on the Independence of Certain Quadratic Forms, Ann. Math. Statist. Vol. 14, 1943.

5) H. Hotelling, Note on a Matric Theorem of A. T. Craig, Ann. Math. Statist. Vol. 15, 1944.

6) H. Sakamoto, On the Independence of Statistics, Kokyuroku of the Institute of Statistical Mathematics, Vol. 1, 1944.

7) K. Matsushita, Note on the Independence of Certain Statistics, Ann. of Instit. Statist. Math. Vol. 1, No. 1, 1949.

8) J. Ogawa, On the Independence of Bilinear and Quadratic Forms of a Random Sample from a Normal Population, Ann. Instit. Statist. Math. Vol. 1, No. 1, 1949.

9) This elegant proof, I owed to Mr. Seiji Nabeya of the Institute of Statistical Mathematics.

10) This was the very point of our earlier papers (7) and (8).

11) See (8).

12) D. S. Villars, A Significance Test and Estimation in the Case of Exponential Regression, Ann. Math. Statist., Vol. 18, 1947.

D. S. Villars & T. W. Anderson, Some Significance Test for Normal Bivariate Distributions, Ann. Math. Statist., Vol. 14 (1943), pp. 141-148.

13) Villars' original criterion was

$$F' = \frac{S_{xx} + 2mS_{xy} + m^2S_{yy}}{m^2S_{xx} - 2mS_{xy} + S_{yy}}$$

where

$$m = \left[S_{yy} - S_{xx} - \sqrt{(S_{yy} - S_{xx})^2 + 4S_{xy}^2} \right] / 2S_{xy},$$

but, I believe, it is more natural for testing the regression to be exponential, that is $m=0$ or ∞ , to adopt the criterion

$$F = \frac{S_{xx} + 2S_{xy} + S_{yy}}{S_{xx} - 2S_{xy} + S_{yy}}.$$

The distribution of F is the so-called F -distribution, when $m=0$ or ∞ .

14) P. C. Tang, The Power Function of the Analysis of Variance Test with Illustrations of their Use, Statist. Res. Mem., Vol. 11.

*) The outline of this note was published in Japanese in Kokyuroku (Research Memoir) of the Institute of Statistical Mathematics, Vol. 5, No. 2, May, 1949.

