ON THE INDEPENDENCE OF CERTAIN ESTIMATES OF VARIANCE[1]

By Allen T. Craig

1. Introduction. It is well known that a necessary and sufficient condition that several statistics be independent in the probability sense, is that the characteristic function of the joint distribution of these statistics shall equal identically the product of the characteristic functions of the distributions of the individual statistics. Thus, if $x_1, x_2, \cdots, x_N$ are N independently observed values of a variable x which is subject to the distribution function $f(x)$, and if $\theta_1, \theta_2, \cdots, \theta_s$ are s statistics, each computed from the N observed values of x, the characteristic function of the joint distribution of the s statistics is given by

$$\varphi(t_1, t_2, \cdots, t_s) = \int \cdots \int e^{it_1\theta_1 + \cdots + it_s\theta_s} f(x_1) \cdots f(x_N)\, dx_N \cdots dx_1.$$

Here, $i = \sqrt{-1}$ and the limits of integration are taken so as to include all admissible values of x. Since the characteristic function of the distribution of θ_v, $v = 1, 2, \cdots, s$, is given by

$$\varphi_v(t_v) = \int \cdots \int e^{it_v\theta_v} f(x_1) \cdots f(x_N)\, dx_N \cdots dx_1,$$

the necessary and sufficient condition for the independence of the s statistics can be written

$$\varphi(t_1, \cdots, t_s) \doteq \varphi_1(t_1) \cdots \varphi_s(t_s), \tag{1}$$

for all real values of $t_1, t_2, \cdots t_s$.

An important phase of sampling theory in statistics is that in which the variable x is subject to the normal distribution function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \qquad -\infty \leq x \leq \infty,$$

and $\theta_1, \cdots, \theta_s$ are s real symmetric quadratic forms in the N independently observed values of x. That is,

$$\theta_1 = \sum_{j=1}^{N} \sum_{k=1}^{N} a_{jk} x_j x_k,$$

$$\theta_2 = \sum_{j=1}^{N} \sum_{k=1}^{N} b_{jk} x_j x_k,$$

$$\vdots$$

$$\theta_s = \sum_{j=1}^{N} \sum_{k=1}^{N} p_{jk} x_j x_k,$$

---

[1] Presented to the Institute of Mathematical Statistics on December 30, 1937, at the invitation of the progam committee. In the paper, we discuss, from a slightly different point of view, some of the material found in the references given at the close of the paper.

so that

$$\varphi(t_1, \cdots, t_s) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{iT}\, dx_N \cdots dx_1; \tag{2}$$

where $T = t_1 \sum \sum a_{jk} x_j x_k + \cdots + t_s \sum \sum p_{jk} x_j x_k - \frac{1}{2i\sigma^2} \sum x_j^2$. If $A_1, \cdots, A_s$ denote the real symmetric matrices of the s quadratic forms, the characteristic function can be written

$$\varphi(t_1, \cdots, t_s) = | I - 2i\sigma^2 t_1 A_1 - \cdots - 2i\sigma^2 t_s A_s |^{-\frac{1}{2}},$$

where I is the unit matrix of order N and the vertical bars indicate the determinant of the matrix within them. Similarly, the characteristic function of the distribution of θ_v is given by

$$\varphi_v(t_v) = | I - 2i\sigma^2 t_v A_v |^{-\frac{1}{2}},$$

so that a necessary and sufficient condition for the independence of the s real symmetric quadratic forms can be written

$$| I - 2i\sigma^2 t_1 A_1 - \cdots - 2i\sigma^2 t_s A_s | = \prod_{v=1}^{s} | I - 2i\sigma^2 t_v A_v |, \tag{3}$$

for all real values of $t_1, t_2, \cdots, t_s$.

Although equation (3) is fundamental and is of considerable value in certain problems, it should be remarked that it is frequently rather tedious to use. This suggests that by strengthening the hypotheses, it may be possible to establish another necessary and sufficient condition which, in certain cases, may be easier to use.

2. **Certain quadratic forms.** In order to lead up to such a theorem as that suggested at the close of the last section, we first consider two theorems regarding real symmetric matrices.

Theorem I. *Let $A_1, A_2, \cdots, A_s$ be s real symmetric matrices, each of order N, such that $A_1 + A_2 + \cdots + A_s = I$, where I is the unit matrix of order N. Let r_v, $v = 1, 2, \cdots, s$, be respectively the ranks of the matrices A_v. If $r_1 + r_2 + \cdots + r_s = N$, each of the non-zero roots of the characteristic equations*[2] *of the matrices A_v is $+1$.*

If $s = 2$, the theorem is almost self-evident. For the characteristic equation of A_2 is $| A_2 - \lambda I | = 0$, which, since $A_1 + A_2 = I$, can be written $| I - A_1 -$

[2] By the characteristic equation of the square matrix A is meant the algebraic equation of degree N in λ. $| A - \lambda I | = 0$. If A is real and symmetric and the rank of A is r, the characteristic equation has exactly r real non-zero roots and $N - r$ zero roots. Cf. Kowalewski, Einführung in die Determinanten-Theorie (1909) pp. 126–128.

$\lambda I \mid = 0$ or $\mid A_1 - (1 - \lambda)I \mid = 0$. But the last equation is the characteristic equation of A_1 with λ replaced by $1 - \lambda$. Thus the roots of the equation $\mid A_1 - \lambda I \mid = 0$ are one minus the roots of $\mid A_2 - \lambda I \mid = 0$. Since the equation $\mid A_2 - \lambda I \mid = 0$ has $N - r_2$ zero roots, the equation $\mid A_1 - \lambda I \mid = 0$ has $N - r_2$ roots equal to $+1$. But $r_1 = N - r_2$ so that all the non-zero roots of $\mid A_1 - \lambda I \mid = 0$ are $+1$. A similar statement holds for the roots of $\mid A_2 - \lambda I \mid = 0$.

In general, we have $A_1 + A_2 + \cdots + A_s = I$ and $r_1 + r_2 + \cdots + r_s = N$. Let $B_1 = A_2 + A_3 + \cdots + A_s$ and denote by R_1 the rank of B_1. Thus[3] $R_1 \leq r_2 + r_3 + \cdots + r_s$. Now $A_1 + B_1 = I$ and the equation $\mid A_1 - \lambda I \mid = 0$ has exactly $N - r_1$ zero roots. Since the roots of $\mid B_1 - \lambda I \mid = 0$ are one minus the roots of $\mid A_1 - \lambda I \mid = 0$, the first of these two equations has at least $N - r_1$ non-zero roots so that $R_1 \geq N - r_1 = r_2 + r_3 + \cdots + r_s$. From $r_2 + r_3 + \cdots + r_s \leq R_1 \leq r_2 + r_3 + \cdots + r_s$ we deduce the equality so that the argument in the case of $s = 2$ applies to the matrices A_1 and B_1. In particular, then, each of the non-zero roots of $\mid A_1 - \lambda I \mid = 0$ is $+1$. By writing $B_2 = A_1 + A_3 + \cdots + A_s$, $B_3 = A_1 + A_2 + A_4 + \cdots + A_s$, and so on, and repeating the argument in each instance, we see that the theorem holds.

Theorem II. *Let $A_1, A_2, \cdots, A_s$ be s real symmetric matrices which satisfy the conditions of Theorem* I. *There then exist $s - 1$ real orthogonal matrices of order N, say $L_1, L_2, \cdots, L_{s-1}$, such that each of the s matrices*

$$L'_{s-1} \cdots L'_1 A_v L_1 \cdots L_{s-1}, \qquad v = 1, 2, \cdots, s,$$

is a diagonal matrix[4] *with the r_v non-zero elements on the principal diagonal equal to $+1$. Necessarily, the sum of these s matrices is the identity matrix.*

In proof of the theorem we shall, to save space, restrict ourselves to the case of $s = 3$, although the method we use will be readily seen to be entirely general. Since A_1 is real and symmetric and since, by Theorem I, the r_1 non-zero roots of the characteristic equation of A_1 are $+1$, there exists a real orthogonal matrix of order N, say L_1, such that

$$L'_1 A_1 L_1 = \left\| \begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \vdots & & \vdots \\ \hline 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \vdots & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \end{array} \right\|$$

where L'_1 is the conjugate of L_1 and where, merely as a convenience of notation, we have placed the r_1 non-vanishing elements of the principal diagonal in the first r_1 rows and columns. If then, in both members of the equation $A_1 + A_2 + A_3 = I$, we multiply on the left by L'_1 and on the right by L_1, we have

[3] Cf. Bôcher, Introduction to Higher Algebra (1921) p. 62.

[4] By a diagonal matrix we mean a matrix whose elements not on the principal diagonal are zero.

$$(4)\qquad \left\|\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \vdots & & \vdots \\ 0 & & & & 0 & & 0 \end{array}\right\| + L_1'A_2L_1 + L_1'A_3L_1 = I,$$

since $L_1'IL_1 = IL_1'L_1 = I$. The matrices $L_1'A_2L_1$ and $L_1'A_3L_1$ are real, symmetric, and the ranks are r_2 and r_3, since L_1 is non-singular. Moreover, the non-zero roots of the characteristic equations of the two matrices are $+1$; for $|L_1'A_2L_1 - \lambda I| = |L_1'(A_2 - \lambda I)L_1| = |L_1'||A_2 - \lambda I||L_1|$, and similarly for the matrix $L_1'A_3L_1$. Now if a real symmetric matrix is positive definite, that is, if all the non-zero roots of its characteristic equation are positive, then[5] all the elements on the principal diagonal are positive or zero, and, if an element on the principal diagonal is zero, all the elements in the row and column in which that element lies are zero. These two facts regarding a real symmetric positive definite matrix, in conjunction with equation (4), require that the matrices $L_1'A_2L_1$ and $L_1'A_3L_1$ be of the forms

$$\left\|\begin{array}{ccc|ccc} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & b_{r_1+1,r_1+1} & \cdots & b_{r_1+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{N,r_1+1} & & b_{NN} \end{array}\right\| \quad\text{and}\quad \left\|\begin{array}{ccc|ccc} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & c_{r_1+1,r_1+1} & \cdots & c_{r_1+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & c_{N,r_1+1} & \cdots & c_{NN} \end{array}\right\|$$

respectively. Now the real symmetric matrix

$$C = \left\|\begin{array}{ccc} b_{r_1+1,r_1+1} & \cdots & b_{r_1+1,N} \\ \vdots & & \vdots \\ b_{N,r_1+1} & \cdots & b_{NN} \end{array}\right\|$$

is of order $N - r_1$, its rank is r_2, and its characteristic equation has r_2 roots equal to $+1$. There then exists a real orthogonal matrix M of order $N - r_1$, say

$$M = \left\|\begin{array}{ccc} m_{r_1+1,r_1+1} & \cdots & m_{r_1+1,N} \\ m_{N,r_1+1} & \cdots & m_{NN} \end{array}\right\|,$$

such that

$$M'CM = \left\|\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \vdots & & \\ \vdots & \vdots & & \vdots & \vdots & & \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{array}\right\|.$$

[5] Cf. Cullis, Matrices and Determinoids (1918) vol. 2, p. 302.

Again, to simplify the notation, we have placed the r_2 non-vanishing elements of the principal diagonal in the first r_2 rows and columns. Consider the orthogonal matrix of order N

$$L_2 = \left\| \begin{array}{ccc|ccc} 1 & 0 \cdots & 0 & 0 & \cdots\cdots & 0 \\ 0 & 1 \cdots & 0 & & & \\ \vdots & & & & & \\ 0 & 0 \cdots & 1 & 0 & \cdots\cdots & 0 \\ \hline 0 & \cdots\cdots & 0 & m_{r_1+1,r_1+1} & \cdots & m_{r_1+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots\cdots & 0 & m_{N,r_1+1} & \cdots\cdots & m_{NN} \end{array} \right\|.$$

It is evident that $L_2'(L_1'A_1L_1)L_2 = L_1'A_1L_1$. If then, both members of $L_1'A_1L_1 + L_1'A_2L_1 + L_1'A_3L_1 = I$ are multiplied on the left by L_2' and on the right by L_2, we get

$$\left\| \begin{array}{ccc|ccc} 1 & 0 \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & & & \\ 0 & \cdots\cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots\cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{array} \right\| + \left\| \begin{array}{ccc|cccc} 0 & \cdots & 0 & 0 & \cdots\cdots & 0 & \cdots 0 \\ \vdots & & \vdots & & & \vdots & \\ 0 & \cdots & 0 & & & 0 & \cdots 0 \\ \hline 0 & \cdots & 0 & 1 & 0 \cdots & 0 & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & 0 & \cdots\cdots & 1 & \vdots \\ \vdots & & \vdots & & & & \vdots \\ 0 & & 0 & \cdots & \cdots & \cdots & 0 \end{array} \right\|$$

$$+ \left\| \begin{array}{ccc|ccc} 0 & \cdots & 0 & 0 & \cdots\cdots & 0 \\ \vdots & & & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots\cdots & 0 \\ \hline 0 & \cdots & 0 & d_{r_1+1,r_1+1} & \cdots & d_{r_1+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & d_{N,r_1+1}, & \cdots & d_{NN} \end{array} \right\| = I.$$

From this last equation, it follows that $d_{jk} = 0, j \neq k, d_{jj} = 0, j = r_1 + 1, \cdots, r_1 + r_2$ and $d_{jj} = 1, j = r_1 + r_2 + 1, \cdots, N$. The third matrix in the left member of preceding equation then takes the form

$$\left\| \begin{array}{ccc|ccc|ccc} 0 & \cdots & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \cdots & 0 & & & & & & \vdots \\ \hline 0 & & & 0 & \cdots & 0 & & & \vdots \\ \vdots & & & \vdots & & \vdots & & & \vdots \\ \vdots & & & 0 & \cdots & 0 & & & 0 \\ \hline \vdots & & & & & & 1 & 0 \cdots & 0 \\ \vdots & & & & & & \vdots & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots\cdots & 1 \end{array} \right\|.$$

This establishes Theorem II when $s = 3$. The procedure may be continued in a fairly obvious manner so as to justify the theorem for any finite positive integer s.

With the aid of Theorems I and II, we are now able to state and prove a very useful theorem on the independence of certain quadratic forms of normally and independently distributed variables. The theorem follows.

Theorem III. *Let* $x_1, x_2, \cdots, x_N$ *be* N *independent values of a normally distributed variable* x *and let* $\theta_1, \cdots, \theta_s$ *be* s *real symmetric quadratic forms in these* N *variables, where* $\sum_1^s \theta_j = \sum_1^N x_j^2$. *If* $r_1, r_2, \cdots, r_s$ *denote respectively the ranks of the quadratic forms, a necessary and sufficient condition that the* s *forms be independent in the probability sense is that* $r_1 + r_2 + \cdots + r_s = N$.

Consider the characteristic function of the joint distribution of the s forms as given by equation (2). In accordance with Theorem II, we can successively introduce new variables by performing real linear transformations with orthogonal matrices $L_1, L_2, \cdots, L_{s-1}$ respectively in such a way that[6] T becomes

$$T = t_1 \sum_1^{r_1} y_j^2 + t_2 \sum_{r_1+1}^{r_1+r_2} y_j^2 + \cdots + t_s \sum_{r_1+\cdots+r_{s-1}+1}^{N} y_j^2 - \frac{1}{2i\sigma^2} \sum_1^N y_j^2.$$

Since each transformation is orthogonal, the absolute value of the Jacobian in each instance is unity. Thus the right member of (2) can now be written as the product of s sets of integrals, the sets containing $r_1, r_2, \cdots, r_s$ integrals respectively. That is,

$$\varphi(t_1, \cdots, t_s) = \varphi_1(t_1) \cdots \varphi_s(t_s),$$

which is equation (1). Hence the theorem.

Under the conditions of Theorem III, the characteristic function of the distribution of θ_v is found by direct integration to be

$$\varphi_v(t_v) = (1 - 2i\sigma^2 t_v)^{-\frac{r_v}{2}}.$$

---

[6] If the variables in a symmetric quadratic form with matrix A are transformed by a linear transformation with matrix B, the new form has the matrix $B'AB$. Cf. Bôcher, p. 129. It should be remarked that these $s - 1$ successive orthogonal transformations can be combined into a single orthogonal transformation with matrix $L = L_1L_2 \cdots L_{s-1}$. For if, by means of a linear transformation with matrix L_1, we pass from the variables $x_1, \cdots, x_N$ to the variables $x_1', \cdots, x_N'$, in which the old variables are expressed explicitly in terms of the new, and thence to variables $x_1'', \cdots, x_N''$ by means of a linear transformation with matrix L_2, the transformation with matrix L_1L_2 will carry us directly from the x's to the x'''s. This extends to any finite number of transformations. Since the product of any two orthogonal matrices is an orthogonal matrix (and hence the product of a finite number of them), we see that the remark is justified. Cf. Bôcher, p. 68 and Kowalewski, p. 161. Note that Bôcher expresses the new variables explicitly in terms of the old.

Thus,

$$f_v(\theta_v)\,d\theta_v = d\theta_v \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-it_v\theta_v}\varphi_v(t_v)\,dt_v$$

$$= \frac{1}{2^{\frac{r_v}{2}}\Gamma\left(\frac{r_v}{2}\right)}\left(\frac{\theta_v}{\sigma^2}\right)^{\frac{r_v}{2}-1} e^{-\frac{1}{2}\left(\frac{\theta_v}{\sigma^2}\right)}\frac{d\theta_v}{\sigma^2},$$

so that the variables θ_v/σ^2 are distributed in accordance with Chi-square distributions with r_v degrees of freedom.[7] Accordingly, when the conditions of Theorem III are satisfied, we may deduce not merely the mutual independence of the θ_v but also the nature of their distributions.

3. **Applications to the analysis of variance.** In the analysis of variance, $N = ab$ independently observed values of a normally distributed variable are classified into a rows and b columns in accordance with some relevant scheme:

$$\begin{matrix} x_{11}, & x_{12}, & \cdots, & x_{1b} \\ x_{21}, & x_{22}, & \cdots, & x_{2b} \\ \vdots & \vdots & & \vdots \\ x_{a1}, & x_{a2}, & \cdots, & x_{ab}. \end{matrix}$$

With the notation $\bar{x}_{j\cdot}$, $\bar{x}_{\cdot k}$, $\bar{x}$ to denote respectively the arithmetic mean of the jth row, the kth column, and the entire set, it is readily seen that

$$(5)\quad \sum_1^a\sum_1^b (x_{jk}-\bar{x})^2 = b\sum_1^a(\bar{x}_{j\cdot}-\bar{x})^2 + a\sum_1^b(\bar{x}_{\cdot k}-\bar{x})^2 + \sum_1^a\sum_1^b (x_{jk}-\bar{x}_{j\cdot}-\bar{x}_{\cdot k}+\bar{x})^2$$

$$= \theta_1 \qquad + \theta_2 \qquad + \theta_3$$

is an identity in the $N = ab$ values of x. It is quite straightforward to exhibit each of the three terms in the right member of (5) as a real symmetric quadratic form in the N variables x_{jk} and to show that the ranks are $r_1 = a - 1$, $r_2 = b - 1$, $r_3 = (a-1)(b-1)$. By the device of adding $\theta_4 = \frac{1}{ab}\left(\sum\sum x_{jk}\right)^2 = N\bar{x}^2$ to both members of (5), we have $\sum\sum x_{jk}^2 = \theta_1 + \theta_2 + \theta_3 + \theta_4$. Moreover, the rank of θ_4 is $r_4 = 1$. Thus $r_1 + r_2 + r_3 + r_4 = ab = N$ and, by Theorem III, we see that the four quadratic forms are mutually independent. In particular, θ_1, θ_2 and θ_3 are independent, and each, measured in units of σ^2, is distributed as is Chi-square with its appropriate number of degrees of freedom.

THE UNIVERSITY OF IOWA.

[7] By the number of degrees of freedom of a real symmetric quadratic form of normally and independently distributed variables, we mean the rank of the matrix of the form.

REFERENCES

(1) M. S. Bartlett and J. Wishart, The distribution of second order moment statistics in a normal system. Proceedings of the Cambridge Philosophical Society, vol. 28 (1931-32) pp. 455-459.

The generalized product moment distribution in a normal system. Same journal, vol. 29 (1932-33) pp. 260-270.

(2) W. G. Cochran, The distribution of quadratic forms in a normal system. Proceedings of the Cambridge Philosophical Society, vol. 30 (1933-34) pp. 178-191.

(3) R. A Fisher, Applications of Student's distribution. Metron, vol. 5 (1925) pp. 90-104.

Statistical Methods for Research Workers (1934) pp. 210-272.

(4) S. S. Wilks, Statistical Inference (1936-37) pp. 38, 44-45.