# BILINEAR FORMS IN NORMALLY CORRELATED VARIABLES

By Allen T. Craig

*University of Iowa*

**1. Summary.** If a variable x is normally distributed with mean zero, we have previously given a necessary and sufficient condition (see references at end of this paper) for the independence of two real symmetric quadratic forms in $n$ independent values of that variable. This condition is that the product of the matrices of the forms should vanish. In the present paper, we have proved that the same algebraic condition is both necessary and sufficient for the independence of two real symmetric bilinear, or a real symmetric bilinear and quadratic form, in normally correlated variables.

**2. Introduction.** In this paper, we determine the moment generating function of the joint distribution of two real symmetric bilinear forms in certain normally correlated variables and derive a necessary and sufficient condition for the independence, in the probability sense, of these forms. We further investigate the condition for independence, in the probability sense, of real symmetric bilinear and quadratic forms.

**3. The moment generating function of the distribution of real symmetric bilinear forms.** Let the two variables $x$ and $y$ have a joint normal distribution with means zero, unit variances and correlation coefficient $\rho$. From this bivariate distribution, repeated random samples of $n$ pairs, say $(x_1, y_1)$, $(x_2, y_2)$, $\cdots$, $(x_n, y_n)$, are drawn. Let $C = || c_{jk} ||$ be a real symmetric matrix and write $\theta = \Sigma\Sigma\, c_{jk}x_jy_k$. The moment generating function of the distribution of $\theta$ is then given by

$$\varphi(t) = E[e^{t\theta}] = \frac{1}{(2\pi\sqrt{1-\rho^2})^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{t\theta - Q}\, dy_n\, dx_n \cdots dy_1\, dx_1,$$

where

$$Q = \frac{1}{2(1-\rho^2)}\sum_j (x_j^2 + y_j^2 - 2\rho x_j y_j)$$

and $\theta$ is defined above. If we subject the $x$'s and $y$'s to the same linear homogeneous transformation with appropriately chosen orthogonal matrix $L$, then $Q$ remains invariant and $\theta$ becomes $\Sigma_j\lambda_j x_j' y_j'$ where the $\lambda$'s are the $n$ real roots of the characteristic equation of $C$, that is, of $| C - \lambda I | = 0$. The integrations are then easily effected and we find that

$$\begin{aligned}\varphi(t) &= \{\prod_j [1 - t(\rho + 1)\lambda_j][1 - t(\rho - 1)\lambda_j]\}^{-\frac{1}{2}},\\ &= \{ | I - t(\rho + 1)C | \cdot | I - t(\rho - 1)C | \}^{-\frac{1}{2}},\\ &= | I - 2\rho t C - (1 - \rho^2)t^2 C^2 |^{-\frac{1}{2}},\end{aligned}$$

where $I$ is the unit matrix of order $n$ and the vertical bars, as usual, indicate the determinant of the enclosed matrix.

Next, let $A = \| a_{jk} \|$ and $B = \| b_{jk} \|$ be two real symmetric matrices each of order $n$. Write $\theta_1 = \Sigma\Sigma a_{jk}x_jy_k$ and $\theta_2 = \Sigma\Sigma b_{jk}x_jy_k$ where the $x$'s and $y$'s are the items of the sample randomly drawn from the bivariate distribution previously described. The moment generating function of the joint distribution of $\theta_1$ and $\theta_2$ is then given by

$$\begin{aligned}\varphi(t_1, t_2) &= E[e^{t_1\theta_1+t_2\theta_2}]\\ &= (2\pi\sqrt{1-\rho^2})^{-n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{t_1\theta_1+t_2\theta_2-Q}\,dy_n\,dx_n\cdots dy_1\,dx_1,\end{aligned}$$

where $\theta_1$, $\theta_2$, and $Q$ have the meanings previously assigned to them. If we pursue a line of reasoning similar to that above, we find that

$$\varphi(t_1, t_2) = \left| I - 2\rho(t_1A + t_2B) - (1-\rho^2)(t_1A + t_2B)^2 \right|^{-\frac{1}{2}}.$$

**4. The independence of bilinear forms.** It is clear that there exist positive numbers, say $h_1$ and $h_2$, such that $\varphi(t_1, t_2)$ exists for $0 < t_1 < h_1$ and $0 < t_2 < h_2$. It is well known that a necessary and sufficient condition for the independence of $\theta_1$ and $\theta_2$ is that $\varphi(t_1, t_2)$ shall factor into the product $\varphi(t_1, 0)\varphi(0, t_2)$. If then, we assume $\theta_1$ and $\theta_2$ to be independent, we have essentially

$$\begin{aligned}&\left| I - 2\rho(t_1A + t_2B) - (1-\rho^2)(t_1A + t_2B)^2 \right|\\ &\quad = \left| I - 2\rho t_1A - (1-\rho^2)t_1^2A^2 \right| \cdot \left| I - 2\rho t_2B - (1-\rho^2)t_2^2B^2 \right|.\end{aligned} \tag{1}$$

If $h$ denotes the smaller of $h_1$ and $h_2$, then the factored form holds for $0 < t_1, t_2 < h$, and hence for all real values of $t_1$ and $t_2$. In particular it holds for $t_2 = t_1$ so that

$$\begin{aligned}&\left| I - 2\rho t_1(A + B) - (1-\rho^2)t_1^2(A + B)^2 \right|\\ &\quad = \left| I - 2\rho t_1A - (1-\rho^2)t_1^2A^2 \right| \cdot \left| I - 2\rho t_1B - (1-\rho^2)t_1^2B^2 \right|.\end{aligned}$$

Let $r_1$, $r_2$, and $r \leq r_1 + r_2$ denote the ranks of the matrices $A$, $B$, and $A + B$. Further let the real non-zero roots of the characteristic equations of these matrices be denoted respectively by $\alpha_1, \alpha_2, \cdots, \alpha_{r_1}$, $\beta_1, \beta_2, \cdots, \beta_{r_2}$, and $\gamma_1, \gamma_2, \cdots, \gamma_r$. Then the members of the preceding equation may be written

$$\prod_{i=1}^{r}[1 - t_1(\rho + 1)\gamma_i][1 - t_1(\rho - 1)\gamma_i]$$

and

$$\prod_{i=1}^{r_1}[1 - t_1(\rho + 1)\alpha_i][1 - t_1(\rho - 1)\alpha_i[\prod_{i=1}^{r_2}[1 - t_1(\rho + 1)\beta_i][1 - t_1(\rho - 1)\beta_i]$$

respectively. It is seen that the left member is a polynomial in $t_1$ of degree $2r$ and that the right member is a polynomial in $t_1$ of degree $2(r_1 + r_2)$. Accordingly, $r = r_1 + r_2$ and the roots $\gamma_1, \cdots, \gamma_1$ consist of the roots $\alpha_1, \cdots, \alpha_{r_1}$, $\beta_1, \cdots, \beta_{r_2}$. That is, if $\theta_1$ and $\theta_2$ are independent, then the rank of $A + B$ is the sum of the ranks of $A$ and $B$ and the non-zero roots of the characteristic equation of $A + B$ consist of those of the characteristic equation of $A$ together with those of $B$. Further, if in (1) we put $t_2 = vt_1$, where $v$ is real, we have

$$\begin{aligned} &| I - 2\rho t_1(A + vB) - (1 - \rho^2)t_1^2(A + vB)^2 | \\ &\quad = | I - 2\rho t_1 A - (1 - \rho^2)t_1^2 A^2 | \cdot | I - 2\rho t_1 vB - (1 - \rho^2)t_1^2 v^2 B^2 |. \end{aligned}$$

Denote the rank of $A + vB$ by $r'$ and the non-zero roots of its characteristic equation by $\delta_1, \cdots, \delta_{r'}$. The immediately preceding equation can then be written

$$\prod_{i=1}^{r'} [1 - t_1(\rho + 1)\delta_i][1 - t_1(\rho - 1)\delta_i]$$

$$= \prod_{i=1}^{r_1} [1 - t_1(\rho + 1)\alpha_i][1 - t_1(\rho - 1)\alpha_i] \prod_{i=1}^{r_2} [1 - t_1(\rho + 1)v\beta_i][1 - t_1(\rho - 1)v\beta_i].$$

From this we infer that, apart from zero roots, the roots of the characteristic equation of $A + vB$ are $\alpha_1, \cdots, \alpha_{r_1}, v\beta_1, \cdots, v\beta_{r_2}$.

If a symmetric matrix, say $M(v)$, has elements which are real polynomials in the real variable $v$, and if the determinant

$$| M(v) - \lambda I | = (-1)^n[\lambda - p_1(v)][\lambda - p_2(v)] \cdots [\lambda - p_n(v)],$$

where $p_1(v), p_2(v), \cdots, p_n(v)$ are likewise real polynomials in $v$, then there exists, for all real values of $v$, a real orthogonal matrix, say $L(v)$, such that

$$L'(v)M(v)L(v) = \left\| \begin{matrix} p_1(v) & 0 & \cdots & 0 \\ 0 & p_2(v) & & \\ \cdot & & \cdot & \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & & & p_n(v) \end{matrix} \right\|.$$

Furthermore[1], $\dfrac{dL(v)}{dv}$ exists for all real values of $v$. Since

$$| A + vB - \lambda I | = (-1)^n \lambda^{n-(r_1+r_2)}(\lambda - \alpha_1) \cdots (\lambda - \alpha_{r_1})(\lambda - v\beta_1) \cdots (\lambda - v\beta_{r_2}),$$

[1] A number of years ago, in connection with another problem, the writer sought the assistance of Professor N. H. McCoy for a proof that $L(v)$ is differentiable at $v = 0$. Professor McCoy's elegant demonstration of the existence of $L(v)$ showed that each element of this orthogonal matrix is itself a real polynomial in $v$, divided by the positive square root of another real polynomial, which polynomial is never negative and which vanishes for no real value of $v$. Thus the derivative of $L(v)$ exists not only for $v = 0$ but for all real values of $v$. The writer thanks Professor McCoy for his kind and generous assistance.

then $A + vB$ belongs to the class $M(v)$ so we have

$$
(2) \qquad L'(v)(A + vB)L(v) = \begin{Vmatrix} \alpha_1 & 0 & \cdots & & & & 0 \\ & \ddots & & & & & \\ 0 & \cdots & \alpha_{r_1} & & \cdots & & 0 \\ 0 & \cdots & & v\beta_1 & \cdots & & 0 \\ \vdots & & & & \ddots & & \\ 0 & & \cdots & & v\beta_{r_2} & \cdots & 0 \\ \vdots & & & & & \ddots & \\ 0 & & & & & & 0 \end{Vmatrix}.
$$

In particular,

$$
(3) \qquad L'(0)AL(0) = \begin{Vmatrix} \alpha_1 & \cdots & & 0 \\ & \ddots & & \\ & & \alpha_{r_1} & \\ & & & 0 \\ \vdots & & & \ddots \\ 0 & & & 0 \end{Vmatrix}.
$$

If we differentiate (2) with respect to $v$ and subsequently set $v = 0$, we have

$$
(4) \qquad \frac{dL'(0)}{dv} AL(0) + L'(0)BL(0) + L'(0)A\frac{dL(0)}{dv} = \begin{Vmatrix} 0 & & \cdots & & & 0 \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \beta_1 & & \\ & & & & \ddots & \\ & & & & \beta_{r_2} & \\ \vdots & & & & & \ddots \\ 0 & & & & & 0 \end{Vmatrix}.
$$

Since $L(v)$ is orthogonal, then $L'(v)L(v) = I$. Upon differentiating both members with respect to $v$, and subsequently setting $v = 0$, it is seen that $\frac{dL'(0)}{dv} L(0) = -L'(0) \frac{dL(0)}{dv}$ so that $L'(0) \frac{dL(0)}{dv}$ is a skew-symmetric matrix, say $S$. Further

$$\frac{dL'(0)}{dv} = -L'(0) \frac{dL(0)}{dv} L'(0) = -SL'(0), \tag{5}$$

and, by taking conjugates,

$$\frac{dL(0)}{dv} = -L(0) \frac{dL'(0)}{dv} L(0) = L(0)S. \tag{6}$$

If we multiply (5) on the right by AL(0) and (6) on the left by $L'(0)A$, we see that (4) may be written

$$L'(0)BL(0) = \begin{Vmatrix} 0 & \cdots & & & & & 0 \\ \vdots & \ddots & & & & & \vdots \\ & & 0 & & & & \vdots \\ & & & \beta_1 & & & \vdots \\ & & & & \ddots & & \vdots \\ & & & & & \beta_{r_2} & \vdots \\ & & & & & & 0 \\ \vdots & & & & & & \ddots \\ 0 & & & & & & 0 \end{Vmatrix} + SL'(0)AL(0) - L'(0)AL(0)S. \tag{7}$$

Since $S$ is skew-symmetric and since $L'(0)AL(0)$ is given by (3), then each element on the principal diagonal of $SL'(0)AL(0)$ and $L'(0)AL(0)S$ is zero. Further, since $L'(0)BL(0)$ is symmetric, then $L'(0)BL(0)$ takes the form

$$\begin{Vmatrix} 0 & k_{12} & \cdots & & & k_{1n} \\ k_{12} & 0 & & & & \\ \vdots & & \ddots & & & \\ & & & \beta_1 & & \\ & & & & \ddots & \\ & & & & \beta_{r_2} & \\ \vdots & & & & & \ddots \\ k_{1n} & & & & & 0 \end{Vmatrix}.$$

Because the non-zero roots of the characteristic equation of $L'(0)BL(0)$ are $\beta_1, \cdots, \beta_{r_2}$ then the sum of all two-rowed principal minors of the determinant of $L'(0)BL(0)$ must equal the sum of the products of $\beta_1, \cdots, \beta_{r_2}$ taken two at a time. That is

$$\sum_{i<j} \beta_i \beta_j = \sum_{i<j} \beta_i \beta_j - \Sigma k_{ij}^2,$$

so that each $k_{ij}$, being real, is zero. Accordingly, $SL'(0)AL(0) - L'(0)AL(0)S$ is a zero matrix and $L'(0)BL(0)$ is given by the first term in the right member of (7). We then have

$$L'(0)AL(0)L'(0)BL(0) = L'(0)ABL(0) = 0,$$

from which it follows that $AB = 0$. Thus, if the real symmetric bilinear forms $\theta_1$ and $\theta_2$ are independent in the probability sense, the product of their matrices is zero.

If, conversely, $AB = 0$, then

$$\begin{aligned}\varphi(t_1, t_2) &= | I - 2\rho(t_1 A + t_2 B) - (1 - \rho^2)(t_1^2 A^2 + t_2^2 B^2) |^{-\frac{1}{2}}, \\ &= | [I - 2\rho t_1 A - (1 - \rho^2) t_1^2 A^2][I - 2\rho t_2 B - (1 - \rho^2) t_2^2 B^2] |^{-\frac{1}{2}}, \\ &= \varphi(t_1, 0)\varphi(0, t_2),\end{aligned}$$

and $\theta_1$ and $\theta_2$ are independent. This establishes the following theorem.

THEOREM I. *Let $x$ and $y$ be normally correlated with means zero, unit variances, and correlation coefficient $\rho$. Let $\theta_1$ and $\theta_2$ be two real symmetric bilinear forms in $n$ random pairs of values of $x$ and $y$, say $(x_1, y_1), (x_2, y_2), \cdots (x_n, y_n)$. A necessary and sufficient condition that $\theta_1$ and $\theta_2$ be independent in the probability sense is that the product of the matrices of the forms be zero.*

**5. Simultaneous reduction of quadratic or bilinear forms.** The argument of Section 4 may be used to establish in a very simple manner the following theorem.

THEOREM II. *Let $A$ and $B$ be two real symmetric matrices with constant elements, each matrix of order $n$. A necessary and sufficient condition that there exist a real orthogonal matrix of order $n$ such that simultaneously each of $L'AL$ and $L'BL$ is in canonical form, wherein no non-zero elements occupy corresponding positions on the principal diagonals, is that $AB = 0$.*

For if such an orthogonal matrix $L$ exists, it is evident that $L'ALL'BL = L'ABL = 0$ from which it follows that $AB = 0$. Conversely, if $AB = 0$, then $v$ being a real scalar, the matrix $(A - \lambda I)(vB - \lambda I)$ is equal to the matrix $-\lambda[(A + vB) - \lambda I]$. These matrices being equal, their determinants are equal so that $A + vB$ belongs to the class $M(v)$ of section 4. Thus $L$ may be taken as $L(0)$ and simultaneously $L'AL$ and $L'BL$ are of the form stated in the theorem.

**6. Independence of bilinaer and quadratic forms.** Let $\theta = \Sigma\Sigma a_{jk} x_j y_k$ be a real symmetric bilinear form of rank $r_1$ in the previously defined variables

$(x_1, y_1), \cdots, (x_n, y_n)$ and let $q = \Sigma\Sigma b_{jk}x_jx_k$ be a real symmetric quadratic form of rank $r_2$ in $x_1, x_2, \cdots, x_n$. As usual, denote the non-zero roots of the characteristic equations of $A$ and $B$ by $\alpha_1, \alpha_2, \cdots, \alpha_{r_1}$ and $\beta_1, \beta_2, \cdots, \beta_{r_2}$ respectively. The moment generating function of the joint distribution of $\theta$ and $q$ is

$$\varphi(t_1, t_2) = \frac{1}{(2\pi\sqrt{1-\rho^2})^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{t_1\theta+t_2q-Q}\, dy_n\, dx_n \cdots dy_1\, dx_1,$$

where, as previously,

$$Q = \frac{1}{2(1-\rho^2)}\Sigma(x_j^2 + y_j^2 - 2\rho x_jy_j).$$

We first orthogonally transform the variables so that the exponent in the integrand becomes, upon writing $\| f_{jk} \| = L'BL$,

$$t_1\Sigma\alpha_j x_j' y_j' + t_2\Sigma\Sigma f_{jk}x_j'x_k' - \frac{1}{2(1-\rho^2)}\Sigma(x_j'^2 + y_j'^2 - 2\rho x_j'y_j').$$

We then integrate on $y_1', y_2', \cdots, y_n'$ and obtain for the exponent in the integrand

$$t_2\Sigma\Sigma f_{jk}x_j'x_k' - \tfrac{1}{2}\Sigma x_j'^2 + \rho t_1\Sigma\alpha_j x_j'^2 + \frac{1-\rho^2}{2} t_1^2\Sigma\alpha_j^2x_j'^2 .$$

If we effect on the variables $x_1', x_2', \cdots, x_n'$ the inverse of the orthogonal transformation initially used on the $x$'s and $y$'s, the exponent in the integrand becomes, using $\| g_{jk} \| = A^2$,

$$t_2\Sigma\Sigma b_{jk}x_jx_k - \tfrac{1}{2}\Sigma x_j^2 + \rho t_1\Sigma\Sigma a_{jk}x_jx_k + \frac{1-\rho^2}{2} t_1^2\Sigma\Sigma g_{jk}x_jx_k$$

or

$$-\tfrac{1}{2}\Sigma\Sigma[\delta_{jk} - 2\rho t_1 a_{jk} - (1-\rho^2)t_1^2 g_{jk} - 2t_2 b_{jk}]x_jx_k ,$$

where $\delta_{jk}$ equals 1 or 0 according as $j$ does or does not equal $k$. Hence,

$$(8) \qquad \varphi(t_1, t_2) = | I - 2\rho t_1 A - (1-\rho^2)t_1^2A^2 - 2t_2B |^{-\frac{1}{2}}.$$

If $\theta$ and $q$ are independent, we have

$$(9) \quad | I - 2\rho t_1 A - (1-\rho^2)t_1^2A^2 - 2t_2B | = | I - 2\rho t_1A - (1-\rho^2)t_1^2A^2 | \cdot | I - 2t_2B |,$$

for $0 < t_1 < h_1$ and $0 < t_2 < h_2$. As before, the members of (9) are polynomials which, being equal for $0 < t_1, t_2 < h$, are equal for all real values of $t_1$ and $t_2$. If we put $t_1 = 1$ and $t_2 = vt_1 = v$, where $v$ is real, then (9) becomes

$$| I - 2\rho A - (1-\rho^2)A^2 - 2vB | = | I - 2\rho A - (1-\rho^2)A^2 | \cdot | I - 2vB |$$
$$= \prod_1^{r_1}[1 - (\rho - 1)\alpha_j][1 - (\rho + 1)\alpha_j]\prod_1^{r_2}(1 - 2v\beta_j).$$

That is,

$$| 2\rho A + (1 - \rho^2)A^2 + 2vB - \lambda I |$$
$$= (-1)^n \lambda^{n-(r_1+r_2)} [\lambda - 2\rho\alpha_1 - (1 - \rho^2)\alpha_1^2] \cdots [\lambda - 2\rho\alpha_{r_1} - (1 - \rho^2)\alpha_{r_1}^2]$$
$$\cdot [\lambda - 2v\beta_1] \cdots [\lambda - 2v\beta_{r_2}]$$

so that $2\rho A + (1 - \rho^2)A^2 + 2vB$ is a matrix of the class $M(v)$. Hence we write

$$L'(v)[2\rho A + (1 - \rho^2)A^2 + 2vB]L(v) = \left\| \begin{matrix} 2\rho\alpha_1 + (1 - \rho^2)\alpha_1^2 & & & & & & & \cdots & 0 \\ & \ddots & & & & & & & \\ & & 2\rho\alpha_{r_1} + (1 - \rho^2)\alpha_{r_1}^2 & & & & & & \\ & & & 2v\beta_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 2v\beta_{r_2} & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 0 \end{matrix} \right\| .$$

The argument of section 4 shows that $L'(0)[2\rho A + (1 - \rho^2)A^2]L(0)L'(0)2BL(0)$ is a zero matrix, from which it follows that $2\rho AB + (1 - \rho^2) A^2B = 0$. But this imposes on $\rho$, $n^2$ conditions of the form

$$2\rho l_{jk} + (1 - \rho^2)m_{jk} = 0, \qquad (j, k = 1, 2, \cdots, n).$$

Since these hold for every $-1 < \rho < 1$, they hold identically. Hence each $l_{jk}$ and $m_{jk}$ is zero. In particular, $\| l_{jk} \| = AB = 0$ if $\theta$ and $q$ are independent.

Conversely, if $AB = 0$, we see by Theorem II that (8) becomes

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

so that $\theta$ and $q$ are independent. This yields Theorem III.

THEOREM III. *Let $x$ and $y$ be normally correlated with means zero, unit variances, and correlation coefficient $\rho$. Let $\theta$ be a real symmetric bilinear form in the $n$ random pairs of values of $x$ and $y$, say $(x_1, y_1), \cdots, (x_n, y_n)$, and let $q$ be a real symmetric quadratic form in $x_1, x_2, \cdots, x_n$ (or $y_1, \cdots, y_n$). A necessary and sufficient condition that $\theta$ and $q$ be independent in the probability sense is that the product of the matrices of the forms be zero.*

For example, let $\theta$ be $n$ times the sample covariance and let $q$ be $n$ times the square of the mean of the $x$'s. Then

$$\theta = \Sigma(x_j - \bar{x})(y_j - \bar{y})$$
$$= \Sigma\Sigma a_{jk}x_j y_k ;$$

where

$$a_{jk} = \frac{n-1}{n} \quad \text{if} \quad j = k,$$

$$= -\frac{1}{n} \text{ otherwise,}$$

and

$$q = n\bar{x}^2 = \Sigma\Sigma b_{jk}x_jx_k, \qquad b_{jk} = 1/n \text{ for } j, k = 1, 2, \cdots, n.$$

Since $AB = 0$, then $\theta$ and $q$ are independent, a fact otherwise known but perhaps not so easily established.

## REFERENCES

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

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