ON THE INDEPENDENCE OF CERTAIN WISHART VARIABLES

BY ROBERT V. HOGG

University of Iowa

1. Introduction and summary. It is well known, in testing an hypothesis concerning the means of several independent normal distributions with common but unknown variance σ^2 , that the likelihood ratio λ , raised to an appropriate positive power, is equal to the ratio of two quadratic forms. That is, there exists a positive constant c so that $\lambda^c = \mathbf{X}'\mathbf{A}\mathbf{X}/\mathbf{X}'\mathbf{B}\mathbf{X}$, where A and B are real symmetric matrices and X is a column matrix whose elements have independent normal distributions. Since $\lambda \leq 1$, we see that $\mathbf{X}'(\mathbf{B} - \mathbf{A})\mathbf{X}$ is a non-negative quadratic form. In addition, we usually find that $X'AX/\sigma^2$ and $X'BX/\sigma^2$ have chi-square distributions. That is, $A^2 = A$ and $B^2 = B$. Consequently, in accordance with a theorem of Hogg and Craig [5], A(B - A) = 0 and $(B - A)^2 =$ (B - A). Hence X'AX and X'(B - A)X are stochastically independent and $X'(B - A)X/\sigma^2$ is chi-square. Hence λ^c has a beta distribution provided each chi-square is central; this is usually the case under the null hypothesis. The analogous situation in multivariate statistical analysis introduces a rather intriguing theorem which almost seems obvious upon first inspection. We will describe this situation after we present some notation and certain preliminary results.

Let the $(n \times p)$ matrix $\mathbf{X} = (x_{ij})$ have the p.d.f.

$$\exp\{-\frac{1}{2}\text{tr}[\mathbf{K}^{-1}(\mathbf{X} - \mathbf{y})'\mathbf{V}^{-1}(\mathbf{X} - \mathbf{y})]\}/(2\pi)^{np/2}|\mathbf{K}|^{n/2}|\mathbf{V}|^{p/2}, \quad - \infty < x_{ij} < \infty,$$

where the $(n \times n)$ matrix V and the $(p \times p)$ matrix K are real symmetric positive definite matrices and \mathbf{v} is a real $(n \times p)$ matrix. If V = I, the $(p \times 1)$ column matrices Y_1, \dots, Y_n of X' are independent and have, respectively, the p-variate normal distributions $N(\mathbf{v}_k, \mathbf{K})$, $k = 1, 2, \dots, n$, where $\mathbf{v}_1, \dots, \mathbf{v}_n$ are the $(p \times 1)$ column matrices of \mathbf{v}' . In a recent article [7], Roy and Gnanadesikan proved the following results which are generalizations of two theorems on quadratic forms [3], [4]. Let A, A_1 , and A_2 be real symmetric $(n \times n)$ matrices. Then X'AX has the Wishart distribution $W(K, r, \mathbf{v}'A\mathbf{v})$ if and only if AVA = A (or $A^2 = A$ provided V = I). Here r is the rank of A and $\mathbf{v}'A\mathbf{v}$ is the matrix of the non-centrality parameters. The forms $X'A_1X$ and $X'A_2X$ are stochastically independent if and only if $A_1VA_2 = 0$ (or $A_1A_2 = 0$ provided V = I). These results permit us to state, without proof, the chi-square decomposition theorem of Hogg and Craig [5] in terms of Wishart variables.

THEOREM 1. Let A, A_1 , \cdots , A_{k-1} , A_k be real symmetric $(n \times n)$ matrices so that $A = A_1 + \cdots + A_{k-1} + A_k$. Let X'AX, $X'A_1X$, \cdots , $X'A_{k-1}X$ have Wishart distributions and let A_k be positive semidefinite. Then $X'A_1X$, \cdots , $X'A_{k-1}X$, $X'A_kX$ are mutually stochastically independent and $X'A_kX$ has a Wishart distri-

Received May 29, 1962; revised March 25, 1963.

935

bution. Moreover, if $A = V^{-1}$ and if k = 2, the above conclusion is still valid even though the hypothesis that A_2 is positive semidefinite is omitted.

Now, in most tests of an hypothesis concerning the means of a multivariate normal distribution with unknown matrix K, the likelihood ratio, raised to an appropriate positive power, is equal to the ratio of two determinants, say $U = |\mathbf{X}'\mathbf{A}\mathbf{X}|/|\mathbf{X}'\mathbf{B}\mathbf{X}|$, where \mathbf{A} and \mathbf{B} are real symmetric matrices with ranks greater than or equal to p. Usually $\mathbf{V} = \mathbf{I}$ and hence, for simplicity, we assume that this is the case. In addition, we frequently know, or it can be easily shown, that both $\mathbf{X}'\mathbf{A}\mathbf{X}$ and $\mathbf{X}'\mathbf{B}\mathbf{X}$ have Wishart distributions. Accordingly, if the fact that the likelihood ratio is less than or equal to one (or $|\mathbf{X}'\mathbf{A}\mathbf{X}| \leq |\mathbf{X}'\mathbf{B}\mathbf{X}|$ for all \mathbf{X}) implies that $\mathbf{B} - \mathbf{A}$ is positive semidefinite, then Theorem 1 requires that $\mathbf{X}'\mathbf{A}\mathbf{X}$ and $\mathbf{X}'(\mathbf{B} - \mathbf{A})\mathbf{X}$ be stochastically independent and that $\mathbf{X}'(\mathbf{B} - \mathbf{A})\mathbf{X}$ have a Wishart distribution. Thus, if this is true, U has a well known distribution, for it is distributed like $|\mathbf{W}_1|/|\mathbf{W}_1 + \mathbf{W}_2|$, where \mathbf{W}_1 and \mathbf{W}_2 are independent Wishart variables. In the next section, we show that this is, in fact, the situation.

Finally, in the last section of this paper, we consider certain other theorems on independence that involve Wishart variables.

2. The theorem. The following theorem, which the author thought was obvious on first observation, is a rather tantalizing one and will no doubt generate a number of alternative proofs. (Since I first constructed a proof of the theorem, with the help of some suggestions of Professor H. T. Muhly, Professor W. T. Reid and a Referee have presented additional proofs. I wish to thank all three of these persons for their interest. The proof in this paper is actually a slight modification of my original one; of the three proofs, it seems to be the one that uses the most elementary concepts of vector spaces and matrix algebra.)

THEOREM 2. Let x be a real $(n \times p)$ matrix. Let A and B be two real symmetric $(n \times n)$ matrices having ranks r and s, respectively, where $p \le r < s$. If $A^2 = A$ and $B^2 = B$, then a necessary and sufficient condition that B - A be positive semi-definite is that $|x'Ax| \le |x'Bx|$ for all real x.

PROOF. First, we prove that $|\mathbf{x}'\mathbf{B}\mathbf{x}| \leq |\mathbf{x}'\mathbf{x}|$ for all \mathbf{x} . Since $\mathbf{B} = \mathbf{B}^2$, there exists an orthogonal matrix \mathbf{L} such that

$$L'BL \,=\, \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}\!.$$

Let y be an $(n \times p)$ matrix defined by $\mathbf{x} = \mathbf{L}\mathbf{y}$. Then $|\mathbf{x}'\mathbf{x}| = |\mathbf{y}'\mathbf{y}|$ and $|\mathbf{x}'\mathbf{B}\mathbf{x}| = |\mathbf{y}'\mathbf{L}'\mathbf{B}\mathbf{L}\mathbf{y}| = |\mathbf{y}'\mathbf{U}'\mathbf{B}\mathbf{L}\mathbf{y}| = |\mathbf{y}'\mathbf{y}|$, where $\mathbf{L}'\mathbf{B}\mathbf{L}\mathbf{y} = \mathbf{y}_0$. Thus, the first s rows of \mathbf{y}_0 are the same as those of \mathbf{y} , but the last n - s rows consist entirely of zero elements. It is known (for example see reference [6], p. 101) that $|\mathbf{y}'\mathbf{y}|$ is equal to the sum of the squares of all $(p \times p)$ determinants formed from the rows of \mathbf{y} . Likewise, $|\mathbf{y}'_0\mathbf{y}_0|$ is a sum of squares; however each term in this latter sum is a term in $|\mathbf{y}'\mathbf{y}|$ and hence $|\mathbf{y}'_0\mathbf{y}_0| \le |\mathbf{y}'\mathbf{y}|$ and $|\mathbf{x}'\mathbf{B}\mathbf{x}| \le |\mathbf{x}'\mathbf{x}|$, for all \mathbf{y} and thus all \mathbf{x} .

Now suppose $|\mathbf{x}'\mathbf{x}| = |\mathbf{x}'\mathbf{B}\mathbf{x}| \neq 0$ and thus $|\mathbf{y}'\mathbf{y}| = |\mathbf{y}_0'\mathbf{y}_0| \neq 0$. We wish to show that $\mathbf{y} = \mathbf{y}_0$ (or $\mathbf{x} = \mathbf{B}\mathbf{x}$). There is at least one $(p \times p)$ determinant of \mathbf{y}_0 that is

non-zero, say the one formed by the first p rows, y_1 , y_2 , \cdots , y_p . Consider each of the last n-s rows of y; let us take y_{s+1} for illustration. Since $|y'y|=|y'_0y_0|$, we have that the determinants

$$\left| egin{array}{c} \mathbf{y}_2 \\ \mathbf{y}_3 \\ \vdots \\ \mathbf{y}_p \\ \mathbf{y}_{s+1} \end{array} \right| = \left| egin{array}{c} \mathbf{y}_1 \\ \mathbf{y}_3 \\ \vdots \\ \mathbf{y}_p \\ \mathbf{y}_{s+1} \end{array} \right| = \cdots = \left| egin{array}{c} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{p-1} \\ \mathbf{y}_{s+1} \end{array} \right| = 0.$$

That is, the vector \mathbf{y}_{s+1} belongs to subspace spanned by the vectors \mathbf{y}_1 , \cdots , \mathbf{y}_{i-1} , \mathbf{y}_{i+1} , \cdots , \mathbf{y}_p , $i=1,2,\cdots$, p. Since \mathbf{y}_1 , \mathbf{y}_2 , \cdots , \mathbf{y}_p are linearly independent, the only vector common to those p subspaces is the zero vector. Accordingly, \mathbf{y}_{s+1} consists only of zero elements. Likewise, \mathbf{y}_{s+2} , \cdots , \mathbf{y}_n are zero vectors and hence $\mathbf{y}=\mathbf{y}_0$. That is, $\mathbf{y}=\mathbf{L}'\mathbf{B}\mathbf{L}\mathbf{y}$ and $\mathbf{L}'\mathbf{x}=\mathbf{L}'\mathbf{B}\mathbf{L}\mathbf{L}'\mathbf{x}$ and $\mathbf{x}=\mathbf{B}\mathbf{x}$ when $|\mathbf{x}'\mathbf{x}|=|\mathbf{x}'\mathbf{B}\mathbf{x}|\neq 0$. Conversely, if $\mathbf{x}=\mathbf{B}\mathbf{x}$, we obviously have $|\mathbf{x}'\mathbf{B}\mathbf{x}|=|\mathbf{x}'\mathbf{B}'\mathbf{B}\mathbf{x}|=|\mathbf{x}'\mathbf{x}|$. So we have proved that if $|\mathbf{x}'\mathbf{B}\mathbf{x}|\neq 0$, then $|\mathbf{x}'\mathbf{B}\mathbf{x}|=|\mathbf{x}'\mathbf{x}|$ if and only if $\mathbf{B}\mathbf{x}=\mathbf{x}$.

Since $|\mathbf{x}'\mathbf{B}\mathbf{x}| \leq |\mathbf{x}'\mathbf{x}|$ for all choices of \mathbf{x} , we know, by choosing \mathbf{x} to be $\mathbf{A}\mathbf{x}$, that $|\mathbf{x}'\mathbf{A}'\mathbf{B}\mathbf{A}\mathbf{x}| \leq |\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x}|$. However the condition of the theorem is $|\mathbf{x}'\mathbf{A}\mathbf{x}| \leq |\mathbf{x}'\mathbf{B}\mathbf{x}|$ for all \mathbf{x} ; thus, by replacing \mathbf{x} by $\mathbf{A}\mathbf{x}$, we see that $|\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x}| \leq |\mathbf{x}'\mathbf{A}'\mathbf{B}\mathbf{A}\mathbf{x}|$. Consequently, it follows that $|(\mathbf{A}\mathbf{x})'\mathbf{A}\mathbf{x}| = |(\mathbf{A}\mathbf{x})'\mathbf{B}(\mathbf{A}\mathbf{x})|$ and thus $\mathbf{A}\mathbf{x} = \mathbf{B}(\mathbf{A}\mathbf{x})$ for all \mathbf{x} such that $|(\mathbf{A}\mathbf{x})'(\mathbf{A}\mathbf{x})| \neq 0$. Since rank $\mathbf{A} = r \geq p$, this requires $\mathbf{A} = \mathbf{B}\mathbf{A} = \mathbf{A}\mathbf{B}$. Finally, we have that $(\mathbf{B} - \mathbf{A})^2 = (\mathbf{B} - \mathbf{A})(\mathbf{B} - \mathbf{A}) = (\mathbf{B} - \mathbf{A})\mathbf{B} = \mathbf{B} - \mathbf{A}$; that is, $\mathbf{B} - \mathbf{A}$ is idempotent and thus positive semi-definite.

Conversely, if $\mathbf{B} - \mathbf{A}$ is positive semidefinite, we have by Theorem 1 that $\mathbf{A}(\mathbf{B} - \mathbf{A}) = \mathbf{0}$ or $\mathbf{A}\mathbf{B} = \mathbf{A}$. But we know, since $\mathbf{A}^2 = \mathbf{A}$, that $|\mathbf{x}'\mathbf{A}\mathbf{x}| \leq |\mathbf{x}'\mathbf{x}|$, for all \mathbf{x} , from an earlier argument. If we choose \mathbf{x} to be $\mathbf{B}\mathbf{x}$, we have that $|\mathbf{x}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{x}| \leq |\mathbf{x}'\mathbf{B}'\mathbf{B}\mathbf{x}|$ and $|\mathbf{x}'(\mathbf{A}\mathbf{B})'(\mathbf{A}\mathbf{B})\mathbf{x}| \leq |\mathbf{x}'\mathbf{B}\mathbf{x}|$ and $|\mathbf{x}'\mathbf{A}\mathbf{x}| \leq |\mathbf{x}'\mathbf{B}\mathbf{x}|$ because $\mathbf{A}\mathbf{B} = \mathbf{A}$. This completes the proof of the theorem.

3. Other results on independence. T. W. Anderson, in his excellent book [2], uses Theorem 4.3.2 (of that book) as a rather fundamental theorem dealing with random variables which have Wishart distributions. We now prove, using Theorem 1, a theorem that is essentially Anderson's Theorem 4.3.2.

THEOREM 3. Let X' be a $(p \times n)$ matrix with independent column matrices which have p-variate normal distributions with positive definite variance-covariance matrix K. Let $E(X') = \Gamma w$, where Γ is a real $(p \times r)$ matrix and w is a real $(r \times n)$ matrix. Let H = ww' be non-singular and let $G = X'w'H^{-1}$. Then GHG' and X'X - GHG' are stochastically independent and have Wishart distributions $W(K, r, \Gamma H\Gamma')$ and W(K, n - r, 0), respectively.

PROOF. Define the matrices A_1 and A_2 by the equations $GHG' = (X'w'H^{-1})H(H^{-1}wX) = X'(w'H^{-1}w)X = X'A_1X$ and $X'X - GHG' = X'X - X'A_1X = X'A_2X$. We see that $A_1^2 = (w'H^{-1}w)(w'H^{-1}w) = w'H^{-1}w = A_1$. Accordingly, the rank of A_1 equals the trace of A_1 which is $tr(w'H^{-1}w) = a_1 + a_2 + a_3 + a_4 + a_4 + a_5 + a_5$

 $\operatorname{tr}(\mathbf{H}^{-1}\mathbf{w}\mathbf{w}') = \operatorname{tr}(\mathbf{I}_r) = r$. Since, in the notation of Section 1, $\mathbf{V} = \mathbf{I}$, $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ is Wishart with non-centrality matrix $(\mathbf{\Gamma}\mathbf{w})\mathbf{A}_1(\mathbf{\Gamma}\mathbf{w})' = \mathbf{\Gamma}\mathbf{w}\mathbf{w}'\mathbf{H}^{-1}\mathbf{w}\mathbf{w}'\mathbf{\Gamma}' = \mathbf{\Gamma}\mathbf{H}\mathbf{\Gamma}'$. However, $\mathbf{X}'\mathbf{X}$ is Wishart with non-centrality matrix $(\mathbf{\Gamma}\mathbf{w})(\mathbf{\Gamma}\mathbf{w})' = \mathbf{\Gamma}\mathbf{H}\mathbf{\Gamma}'$. Thus, by Theorem 1, $\mathbf{X}'\mathbf{A}_1\mathbf{X}$ and $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ are stochastically independent and $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ is $W(\mathbf{K}, n-r, \mathbf{0})$. This completes the proof of the theorem.

In his book, Anderson actually shows that G and $X'A_2X$ are stochastically independent. This raises the more general question: When and only when are the linear forms X'B stochastically independent of the forms X'AX? An answer is provided by the following theorem. The notation of Section 1 is used here.

THEOREM 4. Let **A** be a real symmetric $(n \times n)$ matrix and let **B** be a real $(n \times s)$ matrix. A necessary and sufficient condition that X'AX and X'B be stochastically independent is that AVB = 0 (or AB = 0 if V = I).

PROOF. Since there exists a real symmetric $(n \times n)$ positive definite matrix **D** such that $\mathbf{DD'} = \mathbf{V}$, the transformations $\mathbf{X} = \mathbf{DW}$ and $\mathbf{y} = \mathbf{Dv}$ indicate that we need only to give the proof when $\mathbf{V} = \mathbf{I}$.

Let X_1, \dots, X_p be the columns of X. If X'AX and X'B are independent, then the quadratic form X'_1AX_1 and the s linear forms X'_1B are independent. Thus, by an earlier theorem [1], AB = 0.

Almost as a corollary to Theorem 4, we obtain

THEOREM 5. Let **A** be a real symmetric $(n \times n)$ matrix. If the rank of the real $(n \times s)$ matrix **B** is s, a necessary and sufficient condition that X'AX and X'B be stochastically independent is that X'AX and X'BB'X be stochastically independent.

PROOF. The necessity of the condition is quite obvious. Let us thus consider the proof of the sufficiency of the condition. If X'AX and X'BB'X are independent, we have AVBB' = 0. Hence AVB(B'B) = 0. Since B is of rank s, B'B is a non-singular $(s \times s)$ matrix. Thus AVB = 0; so, by Theorem 4, we have that X'AX and X'B are independent. This completes the proof.

Let us now reconsider Theorem 3. We know (using the notation and results of that theorem) that $GHG' = X'w'H^{-1}wX$ and $X'A_2X$ are independent; so $\mathbf{0} = \mathbf{A}_2(w'H^{-1}w)$, since $\mathbf{V} = \mathbf{I}$ there. Thus $\mathbf{0} = \mathbf{A}_2w'H^{-1}ww'$ or $\mathbf{0} = \mathbf{A}_2w'H^{-1}$ because $\mathbf{w}\mathbf{w}'$ is non-singular. Hence, with Anderson, Theorem 4 requires that $\mathbf{X}'\mathbf{A}_2\mathbf{X}$ and $\mathbf{G} = \mathbf{X}'\mathbf{w}'H^{-1}$ be stochastically independent.

REFERENCES

- [1] AITKEN, A. C. (1950). On the statistical independence of quadratic forms in normal variates. *Biometrika* 37 93-96.
- [2] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, New York.
- [3] CARPENTER, OSMER (1950). Note on the extension of Craig's theorem to non-central variates. Ann. Math. Statist. 21 455-457.
- [4] CRAIG, ALLEN T. (1943). Note on the independence of certain quadratic forms. Ann. Math. Statist. 14 195-197.
- [5] Hogg, Robert V. and Craig, Allen T. (1958). On the decomposition of certain chisquare variables. Ann. Math. Statist. 29 608-610.
- [6] HOHN, FRANZ E. (1958). Elementary Matrix Algebra. Macmillan, New York.
- [7] ROY, S. N. and GNANADESIKAN, R. (1959). Some contributions to ANOVA in one or more dimensions: II. Ann. Math. Statist. 30 318-339.