THE DISTRIBUTIONS OF QUADRATIC FORMS ON NORMAL RANDOM VARIABLES

by

Bruno Baldessari

Department of Statistics
University of North Carolina
and

Università di Roma

Institute of Statistics Mimeo Series No. 476

April 1966

CONTENTS:

This research was supported by the Mathematics Division of The Air Force Office of Scientific Research Contract No. AF-AFOSR-760-65.

DEPARTMENT OF STATISTICS

UNIVERSITY OF NORTH CAROLINA

Chapel Hill, N. C.

1. Introduction and Notation

In this article we obtain a necessary and sufficient condition under which a quadratic form, in normal random variables, is distributed as a given linear combination of independent chi-square variates; in this linear combination the coefficient are arbitrary constants, (Theorem 2). This result is a generalization of the known theorems that a quadratic form, in a set of normal random variables, is distributed as a chi-square variable (or a difference between two independent chi-squares) if and only if the product of the matrix of the quadratic form and the variance-covariance matrix is idempotent (or tripotent).

In addition we give a numerical method for finding the coefficient of this linear combination (section 5). In regard to this type of problem we refer to the text and to the references of [3] and [4] and to the Chapter 4 of [2].

We shall use the following abbreviations: r.v. for random variable; r.vt. for random vector; p.d. for positive definite; s.d. for spectral decomposition; d.f. for degrees of freedom.

X will denote an n-dimensional r.vt. and we assume that X is N(u,V), (i.e., X is distributed like N(u,V) - an n-variate normal distribution with mean vector μ and variance-covariance matrix V, V p.d.). Also we denote by $\chi'^2(n,\lambda)$ a noncentral chi-square r.v. with n d.f. and noncentrality parameter λ, and by S an nxn symmetric matrix, by Λ an nxn diagonal matrix, and by I the nxn identity matrix.

2. A lemma in matrix theory

Lemma 1. <u>If S and V are nxn real symmetric matrices and if V is p.d., then the matrix S V has a spectral decomposition.</u>

<u>Proof:</u> By hypothesis V^{-1} exists and is real, symmetric, p.d., also S is real symmetric, then it exists an nxn real matrix M such that $|M| \neq 0$, $M' V^{-1} M = I$, $M' S M = \Lambda$, where Λ is a real diagonal matrix, and its diagonal elements are the n roots $\lambda_1, \ldots, \lambda_n$ of the equation $|\lambda V^{-1} - S| = 0$, [1].

Let a_j, $j = 1, \ldots, s$, $(s \geq 1)$, be the distinct roots of the equation $|\lambda I - SV| = 0$, (the roots of the equation $|\lambda I - SV| = 0$ are, of course, identical to the roots of the equation $|\lambda V^{-1} - S| = 0$), and let r_j, $j = 1, \ldots, s$, be their respective orders of multiplicity. From this it follows that: $SV = M'^{-1} \Lambda M^{-1} M M' = M'^{-1} \Lambda M'$, and putting $P = M'^{-1}$ we see that: $P^{-1} SVP = \Lambda$. Let B_j, $j = 1, \ldots, s$, be the nxn diagonal matrix shich has elements 1 in the same places in which Λ has element a_j, and 0 otherwise. So we have:

$$P^{-1} S V P = \Lambda = \sum_{j=1}^{s} a_j B_j, \tag{1}$$

and putting

$$E_j = P B_j P^{-1}, \quad j = 1, \ldots, s, \tag{2}$$

we obtain the required spectral decomposition [1]:

$$S V = \sum_{j=1}^{s} a_j E_j \quad . \tag{3}$$

Remark 1. The spectral decomposition (3) is unique except for the fact that, in Λ, we may have $\frac{n!}{r_1!\ldots r_s!}$ different orders of the numbers a_j, $j = 1,\ldots, s$. We consider that one of these orders is chosen and that it is not changed subsequently so that we can consider (3) as the unique s.d. of SV.

Remark 2. The matrices M, P, Λ, B_j, E_j, $j = 1,\ldots,s$, of Lemma 1 will also be found useful in the sequel. We shall not repeat these definitions, it being understood that they are the matrices appearing in the proof of Lemma 1.

Corollary 1. If S and V are nxn real symmetric matrices and if V is p.d., then SV has the s.d.: $S V = \sum_{j=1}^{s} a_j E_j$ where rank $(E_j) = r_j$, $j = 1, \ldots, s$, if and only if:

$$|\lambda I - SV| = \prod_{j=1}^{s} (\lambda - a_j)^{r_j}, \quad a_j \neq a_k, \ j \neq k.$$

Proof: Follows from Lemma 1 and from the definition of s.d. .

3. The moment generating function of the r.v. X' SX.

Lemma 2. If the r. vt. X is N (μ,V) with V p.d., and S is a real symmetric matrix, then the moment generating function $m_{X'SX}(\theta;\mu,V)$ of the r.v. X SX, is:

$$m_{X'SX}(\theta;\mu,V) = |I - 2\theta SV|^{-\frac{1}{2}} \exp\left\{\sum_{j=1}^{s} \frac{\mu' E_j V^{-1}\mu}{2} \frac{2\theta a_j}{1-2\theta a_j}\right\}, \qquad (4)$$

where: a_j, $j = 1, \ldots, s$, are the distinct roots of the equation $|\lambda I - SV| = 0$, (a_j with multiplicity r_j), and where E_j is defined by the s.d.: $SV = \sum_{j=1}^{s} a_j E_j$.

Proof: By definition, we have:

$$m_{X'SX}(\theta;\mu,V) = (2\pi)^{-\frac{n}{2}} |V|^{-\frac{1}{2}} \int \exp\{\theta X' SX - \frac{1}{2}(x^{*}-\mu)' V^{-1}(x-\mu)\} dx$$

where $\int$ is an w fold integral and where $dx \equiv dx_1 \ldots dx_n$. We make the transformation: $x = My + \mu$, so that:

$$m_{X'SX}(\theta;\mu,V) = (2\pi)^{-\frac{n}{2}} \int \exp\{\mu' M'^{-1} + y') \Lambda (y + M^{-1}\mu) - \frac{1}{2} y'y\} dy.$$

Writing: $c' = \mu' M'^{-1} = (c_1, \ldots, c_n)$, say, we have:

$$m_{X'SX}(\theta;\mu,V) = \prod_{i=1}^{n} (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} \exp\{\theta\lambda_i (y_i + c_i)^2 - \frac{1}{2} y_c^2\} dy_i =$$

$$= \left\{\prod_{i=1}^{n} (1 - 2\theta\lambda_i)\right\}^{-\frac{1}{2}} \exp\left\{\sum_{i=0}^{n} \frac{c_i^2}{2} \frac{2\theta\lambda_i}{1-2\theta\lambda_i}\right\}. \qquad (5)$$

As in the proof of Lemma 1, suppose that we have r_j roots λ_i equal to a_j, $j = 1, \ldots, s$, and let: $c_j = \Sigma^{(j)} c_i^2$, where $\Sigma^{(j)}$ is the sum on the r_j

subscript i such that $\lambda_i = a_j$. So we have:

$$m_{X'SX}(\theta;\mu,V)=\prod_{j=1}^{s} (1-2\theta a_j)^{-\frac{r_j}{2}} \left\{\exp \sum_{j=1}^{s} \frac{c_j}{2} \frac{2\theta a_j}{1-2\theta a_j}\right\},$$

where:

$$\prod_{j=1}^{s} (1-2\theta a_j)^{-\frac{r_j}{2}} = |I-2\theta SV|^{-\frac{1}{2}},$$

for:

$$|\lambda I - SV| = \prod_{j=1}^{s} (\lambda - a_j)^{r_j}, \quad \sum_{j=1}^{s} r_j = n,$$

and the equality follows putting $\lambda = (2\theta)^{-1}$ for $\theta \neq 0$; for $\theta = 0$ the equality is verified.

In order to prove the lemma we must show that:

$$\sum_{j=1}^{s} \frac{c_j}{2} \frac{2\theta a_j}{1-2\theta a_j} = \sum_{j=1}^{s} \frac{\mu' E_j V^{-1}\mu}{2} \frac{2\theta a_j}{1-2\theta a_j}, \tag{6}$$

where the matrices E_j are defined by the s.d. $SV = \sum_{j=1}^{s} a_j E_j$, by means of the matrix M.

Decomposing the matrix M^{-1} into its row vectors $m'_1, \ldots, m'_n$ it follows from $c = M^{-1}\mu$ that $c_i = m_i'\mu$. We will denote by $B^{(i)}$ the diagonal matrix which has all elements equal 0 except for the ith element of the diagonal which has the value 1.

By calculation it can be verified that: $c_i^2 = \mu' m_i m_i' \mu = \mu' M'^{-1} B^{(i)} M^{-1}\mu$. So we have: $c_j = \Sigma^{(j)} c_i^2 = \Sigma^{(j)} \mu' m_i m_i' \mu = \mu' M'^{-1} \Sigma^{(j)} B^{(i)} M^{-1}\mu = \mu' M'^{-1} B_j M^{-1}\mu$, $j=1, \ldots, s$. From (2) we have: $E_j = M'^{-1} B_j M'$ so that: $M'^{-1} B_j M^{-1} =$

$M'^{-1} B_j M' M'^{-1} M^{-1} = E_j V^{-1}$, $j = 1, \ldots, s$, therefore: $c_j = \mu' E_j V^{-1} \mu$, $j=1, \ldots, s$, which proves (6) and so also Lemma 2.

Remark 3. The matrices $E_j V^{-1}$, $j = 1, \ldots, s$, are symmetric and positive semidefinite of rank r_j. This follows from the proof of Lemma 2, but it may be also verified as follows: $E_j V^{-1} = M'^{-1} B_j M^{-1}$ and so the matrix $E_j V^{-1}$ is symmetric; the matrix B_j is positive semidefinite and so it exists a real matrix C_j, such that: $B_j = C_j' C_j$, rank (C_j) = rank $(B_j) = r_j$, and therefore: $E_j V^{-1} = M'^{-1} B_j M^{-1} = M'^{-1} C' C M^{-1} = (C M^{-1})' (C M^{-1})$, which is a positive semidefinite matrix of rank equal to the rank of C.

4. The distribution of the r.v. X'SX and the matrix SV.

Theorem 1. If the r. vt. X is $N(\mu,V)$, with V p.d., and if S is a real symmetric matrix, then the r.v. X'SX is distributed as $Y = \sum_{j=1}^{s} a_j \chi_j'^2(r_j; \frac{\mu' L_j \mu}{2})$, $(a_j \neq a_{j'}, j \neq j')$, where the χ'^2's are mutually independent and where: $L_j = L'_j$, $\mu' L_j \mu \geq 0$ rank$(L_j) = r_j$, $j = 1, \ldots, s$, if and only if the matrix SV has the s.d. $SV = \sum_{j=1}^{s} a_j E_j$, with: $\mu' E_j V^{-1}\mu = \mu' L_j \mu$, rank $(E_j) = r_j$ $j = 1, \ldots, s$, $\sum_{j=1}^{s} r_j = n$.

Proof of sufficiency: If SV has the s.d. $SV = \sum_{j=1}^{s} a_j E_j$, with rank $(E_j) = r_j$, $j = 1,\ldots, s$, $\sum_{j=1}^{s} r_j = n$, then, by Corollary 1, we have: $|\lambda I - SV| = \prod_{j=1}^{s} (\lambda - a_j)^{r_j}$, and therefore, putting $\lambda = \frac{1}{2\theta}$, we have:

$$|I - 2\theta SV|^{-\frac{1}{2}} = \prod_{j=1}^{s} (1 - 2\theta a_j)^{-\frac{r_j}{2}} \quad . \tag{7}$$

Also, by Lemma 2, we have:

$$m_{X'SX}(\theta;\mu,V) = |I-2\theta SV|^{-\frac{1}{2}} \exp\left\{ \sum_{j=1}^{s} \frac{\mu' E_j V^{-1}\mu}{2} \frac{2\theta a_j}{1-2\theta a_j} \right\}, \tag{8}$$

and so, by (7), and by the hypothesis $\mu' E_j V^{-1}\mu = \mu' L_j \mu$, $j = 1, \ldots, s$, we see that the r.v. X'SX is distributed like the r.v. Y.

Proof of necessity: If the r.v. X'SX is distributed like the r.v. Y, then:

$$m_{X'SX}(\theta;\mu,V) = \prod_{j=1}^{s} (1-2\theta a_j)^{-\frac{r_j}{2}} \exp\left\{\sum_{j=1}^{s} \frac{\mu' L_j \mu}{2} \; \frac{2\theta a_j}{1-2\theta a_j}\right\}. \tag{9}$$

We observe, now, that the s.d. of SV is independent of μ, (being dependent on S and V only, through M) and so the s.d. of SV corresponding to the r.v. $X'SX$ where X is $N(\mu,V)$, is the same as the s.d. of SV corresponding to the r.v. $Z'SZ$ where the r.vt. Z is $N(0,V)$.[1] But: $m_{Z'SZ}(\theta;0,V) = \prod_{j=1}^{s} (1-2\theta a_j)^{-\frac{r_j}{2}} = |I-2\theta SV|^{-\frac{1}{2}}$, in which we take : $2\theta = \frac{1}{\lambda}$ so that: $|\lambda I-SV| = \prod_{j=1}^{s} (\lambda-a_j)^{r_j}$, and therefore, by Corollary 1, we have the s.d.: $SV = \sum_{j=1}^{s} a_j E_j$, rank $(E_j) = r_j$, $j = 1, \ldots, s$, $\sum_{j=1}^{s} r_j = n$. From the first part of the proof of the sufficiency we know that this implies that, (formulae (7) and (8)), :

$$m_{S'SX}(0;\mu,V) = \prod_{j=1}^{s} (1 - 2\theta a_j)^{-\frac{r_j}{2}} \exp\left\{\sum_{j=1}^{s} \frac{\mu' E_j V^{-1}\mu}{2} \; \frac{2\theta a_j}{1-2\theta a_j}\right\},$$

so that, by the hypothesis (9), we must have:

$$\sum_{j=1}^{s} \frac{\mu'(L_j - E_j V^{-1})\mu}{2} \; \frac{2\theta a_j}{1-2\theta a_j} = 0, \tag{10}$$

for all θ sufficiently small: here we can, obviously, suppose that $a_j \neq 0$, $j=1, \ldots, s$. Expanding $(1-2\theta a_j)^{-1}$ in a geometric series it is easily seen

---

1. This could also be demonstrated by noting that if $R(\theta)$, $S_\ell(\theta)$ are ratios of polynomial functions of θ, then $\exp\{R(\theta)\}$ cannot be expanded number of $S_\ell(\theta)$'s (except in trivial case: $R(\theta)$ = constant).

that (10) implies: $\mu' L_j \mu = \mu' E_j V^{-1}_{\mu} j = 1, \ldots, s$. This completes the proof of Theorem 1.

In the following Theorem 2 we will use the notation: $Y = \sum_{j=1}^{s} a_j \chi_j'^2(\cdot,\cdot)$, $(a_j \neq a_{j'}, j \neq j',)$ to mean that the r.v. Y is a linear combination, with coefficients $a_j, (a_j \neq a_{j'}, j \neq j')$, of s mutually independent non-central chi-square r.v.'s, in which we do not specify the individual d.f.'s and non-centrality parameters except for the fact that the d.f.'s are positive and sum to n, and the non-centrality parameters are non negative.

Theorem 2: <u>If the r. vt. X is $N(\mu,V)$, with V p.d., and if S is a real symmetric matrix, then the r.v. X'SX is distributed like the r.v. $Y = \sum_{j=1}^{s} a_j \chi_j'^2(\cdot,\cdot)$,</u> <u>if and only if</u>:

$$\text{a).}\quad \prod_{j=1}^{s} (SV - a_j I) = 0,$$

$$\text{b).}\quad \prod_{\substack{j=1 \\ j\neq k}}^{s} (SV - a_j I) \neq 0, \; k = 1, \ldots, s. \qquad (11)$$

<u>Proof of necessity</u>: First we show that if X'SX is distributed like Y, then the condition (11,a) is satisfied. In fact, by Theorem and Lemma 1, since the r.v. X'SX is distributed like Y the matrix SV has a s.d.: $SV = \sum_{j=1}^{s} a_j E_j$, where rank (E_j), $j = 1, \ldots, s$, is unspecified except for the fact that: rank $(E_j) \geq 1$, $j = 1, \ldots, s$, $\sum_{j=1}^{s}$ rank $(E_j) = n$. This implies that, for every polynomial $p(x)$, we have [1] : $p(SV) = \sum_{j=1}^{s} p(a_j)E_j$, and so, for the polynomial:

$$(x-a_1) \ldots (x-a_n) \text{ we have: } (SV - a_1 I) \ldots (SV - a_s I) =$$

$$(a_1 - a_1) \ldots (a_1 - a_s) E_1 + \ldots + (a_1 - a_1) \ldots (a_s - a_s) E_s = 0$$

and then the condition (11,a) is satisfied.

We prove, now, that if the r.v. X'SX is distributed like the r.v. Y, then the condition (11,b) is satisfied. For if we suppose that within 1,...,s, there exists a member k such that: $\prod_{\substack{j=1 \\ j\neq k}}^{s} (SV-a_j I) = 0$. Then by $SV = \sum_{t=1}^{s} a_t E_t$, $I = \sum_{t=1}^{s} E_t$, we would have:

$$\prod_{\substack{j=1 \\ j\neq k}}^{s} \left\{ \sum_{t=0}^{s} (a_t - a_j) E_t \right\} = 0, \tag{12}$$

so that, multiplying (12) by E_k and recalling that $E_k E_j = 0$, $t \neq k$, we would have: $\prod_{\substack{j=1 \\ j\neq k}}^{s} (a_k - a_j) E_k = 0$, which implies that: either one of the members a_j, $j\neq k$ is equal to a_k(contradicting to the hypothesis: $a_j \neq a_k$, $j\neq k$) or $E_k = 0$ which contradicts the hypothesis that each of the chi-square r.v.'s involved in the r.v. Y have at least one d.f. since this would imply (Theorem 1) that rank $(E_j) = 0$, $j = 1, \ldots, s$.

<u>Proof of sufficiency</u>: First we prove that the condition (11,a) implies that the r.v. X'SX is distributed, <u>at most</u>, like the r.v. $Y = \sum_{j=1}^{s} a_j \chi_j'^2 (\cdot,\cdot)$, $a_j \neq a_{j'}$, $j =/ j'$. By this we mean that the r.v. X'SX is distributed like a r.v. $Z = \sum_{p=1}^{q} a_{j_p} \chi_{j_p}'^2 (\cdot,\cdot)$, where: $1 \leq q \leq s$, and the members $j_1, \ldots, j_q$ constitute any non empty sub-set of the set $\{1,2, \ldots, s\}$.

Now suppose that the condition (11,a) is satisfied and that the r.v. X'SX is distributed like the r.v.: $T = \sum_{t=1}^{k} b_t \chi_t'^2(\cdot,\cdot)$, with $b_t \neq b_e$, $\ell \neq \ell$, and $k > s$. Then, by Theorem 1 and Lemma L, the matrix SV has the s.d.: $SV= \sum_{t=1}^{k} b_t E_t$ where $b_t \neq b_\ell = 0$, $t \neq \ell$. By (11,a) we have:

$$\left(\sum_{t=1}^{k} b_t F_t - a_1 \sum_{t=1}^{k} F_t\right) \ldots \left(\sum_{t=1}^{k} b_1 F_t - a_s \sum_{t=1}^{k} F_t\right) = 0$$

and multiplying by F_ℓ, $\ell = 1, \ldots, k$, we have:

$$(b_\ell - a_1) F_\ell \left(\sum_{t=1}^{k} b_t F_t - a_2 \sum_{t=1}^{k} F_t\right) \ldots \left(\sum_{t=1}^{k} b_t F_t - a_s \sum_{t=1}^{k} F_t\right) = 0, \; \ell = 1, \ldots, k,$$

so that, continuing the multiplication, we would have:

$$\prod_{j=1}^{s} (b_\ell - a_j) \; F = 0, \; \ell = 1, \ldots, k,$$

but this is possible only if either $F_\ell = 0$, which is excluded, or b_ℓ, $\ell=1,\ldots,k$, is equal to one of the members a_j, $(a_j \neq a_{j'}, \; j \neq j')$, in which case the characteristic roots of the equation $|\lambda I - SV| = 0$ would take their values among the members: $a_1, \ldots, a_s$. Therefore, by Theorem 1, the r.v. X'SX is distributed like a linear combination of non-central chi-square r.v.'s, and the coefficients of this linear combination are among the members $a_1, \ldots, a_s$, i.e., the r.v. X'SX is distributed, at most, like the r.v. Y.

Proof of the sufficiency

In fact, (11,a) implies that the r.v. X'SX is distributed, at most, like the r.v. Y, and this fact leaves open two possibilities: either the r.v. X'SX is distributed like the r.v. Y(and in this case sufficiency follows) or the r.v. X'SX is distributed like the r.v. $Z = \sum_{p=1}^{q} a_{j_p} X'^2_{j_p}(\cdot,\cdot) q \leq s-1, a_{j_p} \neq a_{j_{p'}}, \; p \neq p'$, and we have a contradiction. In fact, in the latter case, by Theorem and Lemma 1, we have the s.d.: $SV = \sum_{p=1}^{q} a_{j_p} E_{j_p}$ and this implies that: $(SV - a_s I) \ldots (SV - a_q I) = 0$, $q \leq s-1$, which contradicts, at least, one of the equations (11,b).

5. Numerical Method

Theorem 2 can be used, also, to justify the following numerical method.

Suppose that we know that the r.vt. X is $N(\mu,V)$, with V p.d. and that we want the distribution of the r.v. X'SX, where the matrix S is real symmetric. We know that the r.v. X'SX is a linear combination of independent non-central chi-square r.v.'s.

A numerical method which gives us the number of $\chi'^2(\cdot,\cdot)$, r.v.'s involved and the coefficients of the linear combination is the following: Consider the sequence of equations, where a, b, c,..., are real numbers:

$$\begin{aligned} &(SV - aI) = 0, \\ &(SV - aI)(SV-bI) = 0, \qquad (13) \\ &(SV - aI)(SV - bI)(SV - cI) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

and find the first of the equations (13), which is satisfied by a finite set of real numbers: a, b, c, ..., . Then the r.v. X'SX is distributed like the r.v. $Y = a\,\chi_1'^2(\cdot,\cdot) + b\,\chi_2'^2(\cdot,\cdot) + c\,\chi_3'^2(\cdot,\cdot) + \ldots$

This method is, obviously, justified by the Theorem 2.

If the number of chi-squares involved in the linear combination is reasonably low, then we may obtain the d.f. and the non-centrality parameters by equating the moments of the r.v. X'SX and of the linear combination of chi-squares.

A C K N O W L E D G E M E N T S

The author expresses his gratitude to Professor N.L. Johnson for kind encouragements, helpful discussions, and for revising this article.

REFERENCES

[1] Ayres, F. (1962). Theory and problems in Matrices, Schaum, New York.

[2] Graybill, F. A. (1961). An introduction to linear statistical models, Vol. I, Mc Graw Hill, New York.

[3] Graybill, F. A. and Marsaglia, G. (1957). Independent matrices and quadratic forms in the general linear hypothesis, Ann. Math. Statistics, 28, 678-686.

[4] Luther, N. Y. (1965). Decomposition of symmetric matrices and distributions of quadratic forms, Ann. Math. Statistics, 36, 683-690.