The Annals of Mathematical Statistics
1969, Vol. 40, No. 4, 1430–1438

QUADRATIC FORMS AND IDEMPOTENT MATRICES WITH RANDOM ELEMENTS[1]

BY FRANKLIN A. GRAYBILL AND GEORGE MILLIKEN

Colorado State University

1. Introduction. There have been a number of papers on the distribution of quadratic forms of normal variables [1], [2], [3], [4], [5]. The results are of particular importance in the theory of the general linear model, and idempotent matrices play a significant role in the distribution properties of quadratic forms for these models. In fact there are two basic results: let $\mathbf{y}$ be distributed as an $n \times 1$ normal random vector with mean $\boldsymbol{\mu}$ and positive definite covariance matrix $\mathbf{V}$. (1) $\mathbf{y}'\mathbf{A}\mathbf{y}$ is distributed as a non-central chi-square if and only if $\mathbf{AV}$ is idempotent; (2) $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are independent if and only if $\mathbf{AVB} = \mathbf{0}$.

In the theorems in these papers mentioned above the matrices of the quadratic forms have constant elements. The purpose of this paper is to extend some of these theorems to the case where the elements of the matrices are random variables, and, of course, in these cases the function may no longer be a quadratic form in the observation vector $\mathbf{y}$.

2. Preliminary lemmas. In the theorems that we shall prove the basic variables will be assumed to have a multivariate normal distribution. Therefore, we shall state some results on this distribution.

DEFINITION 2.1. *Multivariate normal distribution of rank k.* Let $\mathbf{y}$ be an $n \times 1$ random vector with distribution function $F_y(\cdot)$ and characteristic function $\phi_y(\cdot)$. The vector $\mathbf{y}$ is defined to have a multivariate normal distribution of rank k if and only if the characteristic function of $\mathbf{y}$ is defined by

$$\phi_y(\mathbf{t}) = \exp\,(i\boldsymbol{\mu}'\mathbf{t} - \tfrac{1}{2}\mathbf{t}'\mathbf{V}\mathbf{t}); \quad \text{for all } \mathbf{t} \text{ in } n\text{-dimensional real space;}$$

where $\mathbf{V}$ is a non-negative (definite) $n \times n$ matrix of rank k and with constant elements, $\boldsymbol{\mu}$ is an $n \times 1$ vector of constant elements and $\boldsymbol{\mu}$ is in the column space of $\mathbf{V}$.

We shall also use the notation $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{V})$, $\mathbf{V}$ is $n \times n$ of rank k; to denote the distribution of $\mathbf{y}$.

We shall state a number of lemmas concerning the multivariate normal that we shall refer to later.

LEMMA 2.1. *Let $\mathbf{y}$ be defined in Definition 2.1. Then $\mathcal{E}(\mathbf{y}) = \boldsymbol{\mu}$; Cov $(\mathbf{y}) = \mathbf{V}$ where $\mathcal{E}(\cdot)$ denotes expectation and Cov $(\cdot)$ denotes a covariance matrix.*

LEMMA 2.2. *Let $\mathbf{y}$ be defined in Definition 2.1. Then there exists an $n \times k$ matrix $\mathbf{H}$ of rank k and a $k \times 1$ vector $\boldsymbol{\theta}$ such that $\mathbf{y} = \mathbf{H}(\mathbf{z} + \boldsymbol{\theta})$ where $\mathbf{z}$ is a $k \times 1$ vector*

Received 18 June 1968.
[1] This work was supported by National Institutes of Health Training Grant No. 5 TI GM–725.

of independent random variables each distributed as the standard univariate normal, i.e. mean zero and variance one. (*In fact* $\mathbf{H}$ *is any* $n \times k$ *matrix of rank* k *such that* $\mathbf{V} = \mathbf{H}\mathbf{H}'$).

LEMMA 2.3. *In Lemma* 2.2 *we can write* $\mathbf{z} = \mathbf{H}^c\mathbf{y} - \boldsymbol{\theta}$ *where* $\mathbf{H}^c$, *called a conditional inverse of* $\mathbf{H}$, *is any matrix* $\mathbf{H}^c$ *such that* $\mathbf{H}\mathbf{H}^c\mathbf{H} = \mathbf{H}$.

LEMMA 2.4. *Let* $\mathbf{y}$ *be defined by Definition* 2.1 *and let* $\mathbf{A}$ *be an* $n \times n$ *matrix with constant elements. The quadratic form* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *is distributed as a non-central chi-square variable with* m *degrees of freedom if and only if* $\mathbf{H}'\mathbf{A}\mathbf{H}$ *is idempotent and* $\operatorname{tr}(\mathbf{H}'\mathbf{A}\mathbf{H}) = m$ *where* $\mathbf{H}$ *is any* $n \times k$ *matrix of rank* k *such that* $\mathbf{V} = \mathbf{H}\mathbf{H}'$. (*The non-centrality parameter is* $\frac{1}{2}\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$).

LEMMA 2.5. *Let* $\mathbf{y}$ *be defined in Definition* 2.1. *The two quadratic forms* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *and* $\mathbf{y}'\mathbf{B}\mathbf{y}$ *are independent if and only if* $\mathbf{H}'\mathbf{A}\mathbf{V}\mathbf{B}\mathbf{H} = \mathbf{0}$ *where* $\mathbf{H}$ *is any* $n \times k$ *matrix such that* $\mathbf{V} = \mathbf{H}\mathbf{H}'$ ($\mathbf{A}$ *and* $\mathbf{B}$ *have constant elements*).

LEMMA 2.6. *Let* $\mathbf{y}$ *be defined in Definition* 2.1. *A sufficient condition for the quadratic form* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *to be distributed as a non-central chi-square with* m *degrees of freedom is that* $\mathbf{A}\mathbf{V}$ *is idempotent of rank* m ($\mathbf{A}$ *has constant elements*).

LEMMA 2.7. *Let* $\mathbf{y}$ *be defined in Definition* 2.1. *A sufficient condition for the two quadratic forms* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *and* $\mathbf{y}'\mathbf{B}\mathbf{y}$ *to be independent is that* $\mathbf{A}\mathbf{V}\mathbf{B} = \mathbf{0}$ ($\mathbf{A}$ *and* $\mathbf{B}$ *have constant elements*).

3. The main theorems. In the previous section the matrices $\mathbf{A}$ and $\mathbf{B}$ of the quadratic forms were assumed to have constant elements. In this section we shall generalize some of the results to include the case when the elements of $\mathbf{A}$ and $\mathbf{B}$ may be functions of $\mathbf{y}$.

THEOREM 3.1. *Let the* $n \times 1$ *random vector* $\mathbf{y}$ *be such that* $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{I})$. *Let* $\mathbf{K}$ *be any non-zero* $r \times n$ *matrix of constants of rank* $k < n$; *let* $\mathbf{L}$ *be any non-zero* $n \times n$ *matrix of constants such that the rows of* $\mathbf{L}$ *are in the orthogonal complement of the row space of* $\mathbf{K}$. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix with elements* a_{ij} *where* $a_{ij} = f_{ij}(\mathbf{K}\mathbf{y})$, *and where* $f_{ij}(\cdot)$ *is a Borel function of the random vector* $\mathbf{K}\mathbf{y}$. *The random variable* $w = \mathbf{y}'\mathbf{A}\mathbf{y}$ *is distributed as a non-central chi-square if the following four conditions hold with probability one.*

(1) $\mathbf{A} = \mathbf{L}'\mathbf{A}\mathbf{L}$;

(2) $\mathbf{A}$ *is idempotent*;

(3) $\operatorname{tr}(\mathbf{A}) = m$; m *is a constant positive integer*;

(4) $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \lambda$; λ *is a constant.*

PROOF. Define the random variable $\boldsymbol{\mu}$ by

$$\mathbf{u} = \begin{bmatrix} \mathbf{K} \\ \mathbf{L} \end{bmatrix} \mathbf{y} = \begin{bmatrix} \mathbf{K}\mathbf{y} \\ \mathbf{L}\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}.$$

Then $\mathbf{u}_1 \sim N(\mathbf{K}\boldsymbol{\mu}, \mathbf{K}\mathbf{K}')$, $\mathbf{u}_2 \sim N(\mathbf{L}\boldsymbol{\mu}, \mathbf{L}\mathbf{L}')$ and $\mathbf{u}_1$ is independent of $\mathbf{u}_2$ since $\mathbf{L}\mathbf{K}' = \mathbf{0}$ (i.e. the rows of $\mathbf{L}$ are in the orthogonal complement of the row space of $\mathbf{K}$). From condition (1) we obtain $w = \mathbf{y}'\mathbf{A}\mathbf{Y} = \mathbf{y}'\mathbf{L}'\mathbf{A}\mathbf{L}\mathbf{y} = \mathbf{u}_2'\mathbf{A}\mathbf{u}_2$. Since $\mathbf{A}$ depends only on the random vector $\mathbf{u}_1$ and since $\mathbf{u}_1$ and $\mathbf{u}_2$ are independent, the distribution of the conditional random variable $w \mid \mathbf{u}_1 = \mathbf{u}_1^*$ is by Lemma 2.4

non-central chi-square with m degrees of freedom if conditions (2), (3) and (4) hold. But this distribution is the same for every allowable value of $\mathbf{u}_1^*$, hence the marginal distribution of w is non-central chi-square with m degrees of freedom.

THEOREM 3.2. *Let* $\mathbf{y}$, $\mathbf{K}$ *and* $\mathbf{L}$ *be defined as in Theorem* 3.1. *Let the elements of the* $n \times n$ *matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *be Borel functions of the vector* $\mathbf{Ky}$. *The two random variables* w_1 *and* w_2, *where* $w_1 = \mathbf{y}'\mathbf{Ay}$ *and* $w_2 = \mathbf{y}'\mathbf{By}$, *are independent if the following nine conditions hold with probability one*:

(1) $\mathbf{L}'\mathbf{AL} = \mathbf{A}$;
(2) $\mathbf{L}'\mathbf{BL} = \mathbf{B}$;
(3) $\mathbf{A} = \mathbf{A}^2$;
(4) $\mathbf{B} = \mathbf{B}^2$;
(5) $\mathrm{tr}(\mathbf{A}) = m_1$;
(6) $\mathrm{tr}(\mathbf{B}) = m_2$;
(7) $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \lambda_1$;
(8) $\boldsymbol{\mu}'\mathbf{B}\boldsymbol{\mu} = \lambda_2$;
(9) $\mathbf{AB} = \mathbf{0}$.

where m_1, m_2 *are constant positive integers*, λ_1 *and* λ_2 *are constants.*

PROOF. If we use the notation we used in the proof of Theorem 3.1, by conditions (1) and (2) we can write

$$w_1 = \mathbf{y}'\mathbf{Ay} = \mathbf{y}'\mathbf{L}'\mathbf{ALy} = \mathbf{u}_2'\mathbf{Au}_2$$
$$w_2 = \mathbf{y}'\mathbf{By} = \mathbf{y}'\mathbf{L}'\mathbf{BLy} = \mathbf{u}_2'\mathbf{Bu}_2.$$

By Lemma 2.2 we can write $\mathbf{u}_1 = \mathbf{H}_1(\mathbf{z}_1 + \mathbf{H}_1^c\mathbf{K}\boldsymbol{\mu})$; $\mathbf{u}_2 = \mathbf{H}_2(\mathbf{z}_2 + \mathbf{H}_2^c\mathbf{L}\boldsymbol{\mu})$ where $\mathbf{H}_1\mathbf{H}_1' = \mathbf{KK}'$; $\mathbf{H}_2\mathbf{H}_2' = \mathbf{LL}'$; $\mathbf{H}_1$ has rank k; $\mathbf{H}_2$ has rank l. Also let $\mathbf{H}_1^c\mathbf{K}\boldsymbol{\mu} = \boldsymbol{\theta}_1$; $\mathbf{H}_2^c\mathbf{L}\boldsymbol{\mu} = \boldsymbol{\theta}_2$. Therefore $\mathbf{z}_1 \sim N(\mathbf{0}, \mathbf{I})$; $\mathbf{z}_2 \sim N(\mathbf{0}, \mathbf{I})$; $\mathbf{z}_1$ and $\mathbf{z}_2$ are independent; $w_1 = (\mathbf{z}_2 + \boldsymbol{\theta}_2)'\mathbf{H}_2'\mathbf{AH}_2(\mathbf{z}_2 + \boldsymbol{\theta}_2)$; $w_2 = (\mathbf{z}_2 + \boldsymbol{\theta}_2)'\mathbf{H}_2'\mathbf{BH}_2(\mathbf{z}_2 + \boldsymbol{\theta}_2)$ where $\mathbf{A}$ and $\mathbf{B}$ are functions of $\mathbf{z}_1$. We shall determine the characteristic function of w_1, w_2:

$$\begin{aligned}\phi_{w_1,w_2}(t_1, t_2) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\,[i(t_1w_1 + t_2w_2)]\, dF(\mathbf{z}_2)\, dF(\mathbf{z}_1) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\,[i(\mathbf{z}_2 + \boldsymbol{\theta}_2)'\mathbf{H}_2'\mathbf{AH}_2(\mathbf{z}_2 + \boldsymbol{\theta}_2)t_1 \\ &\qquad + i(\mathbf{z}_2 + \boldsymbol{\theta}_2)'\mathbf{H}_2'\mathbf{BH}_2(\mathbf{z}_2 + \boldsymbol{\theta}_2)t_2]\, dF(\mathbf{z}_2)\}\, dF(\mathbf{z}_1) \\ &\qquad -\infty < t_1 < \infty;\ -\infty < t_2 < \infty\end{aligned} \tag{2.1}$$

where $dF(\mathbf{z}_1) = (2\pi)^{-k/2} \exp(-\frac{1}{2}\mathbf{z}_1'\mathbf{z}_1')\, d\mathbf{z}_1$; $dF(\mathbf{z}_2) = (2\pi)^{-l/2} \exp(-\frac{1}{2}\mathbf{z}_2'\mathbf{z}_2)\, d\mathbf{z}_2$. The quantity in braces is a constant; i.e. does not depend on $\mathbf{z}_1$. We argue as follows. We chose $\mathbf{H}_2$ such that $\mathbf{LL}' = \mathbf{H}_2\mathbf{H}_2'$. By (1) and (3) it follows that $\mathbf{H}_2'\mathbf{AH}_2$ is idempotent; by (2) and (4) it follows that $\mathbf{H}_2'\mathbf{BH}_2$ is idempotent; by (5) $m_1 = \mathrm{tr}\,(\mathbf{A}) = \mathrm{tr}\,(\mathbf{L}'\mathbf{AL}) = \mathrm{tr}\,(\mathbf{ALL}') = \mathrm{tr}\,(\mathbf{AH}_2\mathbf{H}_2') = \mathrm{tr}\,(\mathbf{H}_2'\mathbf{AH}_2)$; similarly $m_2 = \mathrm{tr}\,(\mathbf{H}_2'\mathbf{BH}_2)$, and hence $\mathrm{tr}\,(\mathbf{H}_2'\mathbf{AH}_2)$ and $\mathrm{tr}\,(\mathbf{H}_2'\mathbf{BH}_2)$ are constant positive integers. By (7) and (1) we obtain $\lambda_1 = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{L}'\mathbf{AL}\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{L}'(\mathbf{H}_2^c)'\mathbf{H}_2'\mathbf{AH}_2\mathbf{H}_2^c\mathbf{L}\boldsymbol{\mu} = \boldsymbol{\theta}_2'\mathbf{H}_2'\mathbf{AH}_2\boldsymbol{\theta}_2$; similarly $\lambda_2 = \boldsymbol{\theta}_2'\mathbf{H}_2'\mathbf{BH}_2\boldsymbol{\theta}_2$. Thus $\boldsymbol{\theta}_2'\mathbf{H}_2'\mathbf{AH}_2\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_2'\mathbf{H}_2'\mathbf{BH}_2\boldsymbol{\theta}_2$ are constants.

By condition (9) it follows that $\mathbf{H}_2'\mathbf{A}\mathbf{H}_2\mathbf{H}_2'\mathbf{B}\mathbf{H}_2 = \mathbf{0}$. The quantity in braces in (2.1) by straightforward evaluation reduces to

$$(2.2)\quad \{(1 - 2it_1)^{-m_1/2} \exp [it_1\lambda_1/(1 - 2it_1)]\} \cdot\{(1 - 2it_2)^{-m_2/2} \exp [it_2\lambda_2/(1 - 2it_2)]\}.$$

But this does not depend on $\mathbf{z}_1$; hence the characteristic function of w_1, w_2 is the quantity in (2.2). From the characteristic function we notice that w_1 and w_2 are independent and that w_i is a non-central chi-square random variable with m_i degrees of freedom and with non-centrality λ_i for $i = 1, 2$. This completes the proof of the theorem.

Next we shall generalize the following theorem. Let the n independent random vectors $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n$ be distributed $N(\mathbf{0}, \mathbf{V})$ where $\mathbf{V}$ is a $p \times p$ matrix of rank p. The matrix $\sum_{j=1}^n \sum_{i=1}^n \mathbf{y}_i\mathbf{y}_j' a_{ij}$ is distributed as $W(p, k, \mathbf{V})$ if $\mathbf{A} = [a_{ij}]$ is an idempotent matrix of constants of rank k where $W(p, k, \mathbf{V})$ denotes a Wishart distribution [1]. We shall generalize this to the case where a_{ij} is a function of the vectors $\mathbf{y}_i$.

THEOREM 3.3. *Suppose that the n vectors* $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n$ *are jointly independent and each is distributed* $N(\mathbf{0}, \mathbf{V})$ *where* $\mathbf{V}$ *is a* $p \times p$ *positive definite matrix. Let* $\mathbf{u}_i$ *and* $\mathbf{w}_i$ *be defined by* $\mathbf{y}_i = \begin{bmatrix} \mathbf{u}_i \\ \mathbf{w}_i \end{bmatrix}$ *where* $\mathbf{u}_i$ *has dimension* $p_1 \times 1$ *where* $0 < p_1 < p$. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix such that each element* a_{ij} *is a Borel function of the vectors* $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n$. *The random matrix* $\mathbf{S} = \sum_{j=1}^n \sum_{i=1}^n \mathbf{u}_i\mathbf{u}_j' a_{ij}$ *is distributed as* $W(p_1, k, \mathbf{V}_{11.2})$ *where* $\mathbf{V}_{11.2} = \mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ *if the following conditions hold with probability one.*

(1) $\mathbf{A} = \mathbf{A}^2$;

(2) tr $(\mathbf{A}) = k$; *k is a constant positive integer*;

(3) $\mathbf{R}_{12}[\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n]\mathbf{A} = \mathbf{0}$.

PROOF. We shall define $\mathbf{R}$ to be $\mathbf{V}^{-1}$ and partition $\mathbf{V}$ and $\mathbf{R}$ so that

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}; \qquad \mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix}$$

where $\mathbf{V}_{11}$ and $\mathbf{R}_{11}$ have dimension $p_1 \times p_1$. We define the $np \times 1$ vector $\mathbf{y}$, the $np_1 \times 1$ vector $\mathbf{u}$ and the $np_2 \times 1$ vector $\mathbf{w}$ by $\mathbf{u}' = [\mathbf{u}_1', \mathbf{u}_2', \cdots, \mathbf{u}_n']$ and $\mathbf{w}' = [\mathbf{w}_1', \mathbf{w}_2', \cdots, \mathbf{w}_n']$, $\mathbf{y}' = [\mathbf{y}_1', \mathbf{y}_2', \cdots, \mathbf{y}_n']$ and it follows that $\mathbf{u} \sim N(\mathbf{0}, \mathbf{V}_{11} \times \mathbf{I})$; $\mathbf{w} \sim N(\mathbf{0}, \mathbf{V}_{22} \times \mathbf{I})$; $\mathbf{y} \sim N(\mathbf{0}, \mathbf{V} \times \mathbf{I})$. (The notation $\mathbf{A} \times \mathbf{B}$ will mean the left direct product of $\mathbf{A}$ and $\mathbf{B}$). We define the $p_1 \times p_1$ symmetric matrix $\mathbf{T}$ to have real elements t_{ii} on the ith diagonal and $\frac{1}{2}t_{ij}$ on the ijth off-diagonal. We define the vector $\mathbf{t}$ by $\mathbf{t}' = (t_{11}, \frac{1}{2}t_{12}, t_{22}, \frac{1}{2}t_{13}, \frac{1}{2}t_{23}, t_{33}, \cdots, t_{p_1p_1})$. We shall find the characteristic function of $\mathbf{S}$. We obtain

$$(3.1)\quad \begin{aligned} \phi_S(\mathbf{t}) &= \mathcal{E} \exp [i \operatorname{tr} (\mathbf{TS})] \\ &= \int_{-\infty}^{\infty} (2\pi)^{-np/2}|\mathbf{V} \times \mathbf{I}|^{-\frac{1}{2}} \exp [i \operatorname{tr} (\mathbf{TS})] \exp [-\tfrac{1}{2}\mathbf{y}'(\mathbf{V} \times \mathbf{I})^{-1}\mathbf{y}]\, d\mathbf{y}. \end{aligned}$$

We shall examine the exponent of the integrand. We obtain

$$-\tfrac{1}{2}[\mathbf{u}'(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})\mathbf{u} + \mathbf{u}'(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w} + \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{I})\mathbf{u} + \mathbf{w}'(\mathbf{R}_{22} \times \mathbf{I})\mathbf{w}]$$
$$= -\tfrac{1}{2}\{[\mathbf{u} - (\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}]'[\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A}]$$
$$\cdot[\mathbf{u} - (\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}]$$
$$+ \mathbf{w}'[(\mathbf{R}_{22} \times \mathbf{I}) - (\mathbf{R}_{21} \times \mathbf{I})(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}(\mathbf{R}_{12} \times \mathbf{I})]\mathbf{w}\}$$
$$= -\tfrac{1}{2}\{q_1 + q_2\}.$$

If we substitute into (3.1) we obtain

$$\phi_S(\mathbf{t}) = (2\pi)^{-np/2}|\mathbf{V} \times \mathbf{I}|^{-\frac{1}{2}} \int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} e^{-q_1}\, d\mathbf{u}]e^{-q_2}\, d\mathbf{w}.$$

The value of the integral in the braces is clearly equal to

$$(2\pi)^{p_1 n/2}|\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A}|^{-\frac{1}{2}} = |\mathbf{R}_{11} - 2i\mathbf{T}|^{-k/2}|\mathbf{R}_{11}|^{-n/2+k/2}(2\pi)^{p_1 n/2}$$

since $\mathbf{A}$ is idempotent with tr $(\mathbf{A}) = k =$ rank $(\mathbf{A})$. Thus we obtain

$$\phi_S(\mathbf{t}) = (2\pi)^{-np_2/2}|\mathbf{R}_{11}|^{-n/2+k/2}|\mathbf{V}|^{-n/2}|\mathbf{R}_{11} - 2i\mathbf{T}|^{-k/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-q_2}\, d\mathbf{w}.$$

To evaluate the integral we write

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-q_2}\, d\mathbf{w}$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\{-\tfrac{1}{2}\mathbf{w}'[(\mathbf{R}_{22} \times \mathbf{I}) \tag{3.2}$$
$$- (\mathbf{R}_{12} \times \mathbf{I})(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}(\mathbf{R}_{12} \times \mathbf{I})]\mathbf{w}\}\, d\mathbf{w}.$$

We can write the exponent as

$$q_2 = \tfrac{1}{2}\mathbf{w}'(\mathbf{R}_{22} \times \mathbf{I})\mathbf{w} - \tfrac{1}{2}\mathbf{w}'[(\mathbf{R}_{12} \times \mathbf{I})(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}(\mathbf{R}_{12} \times \mathbf{I})]\mathbf{w}$$
$$= \tfrac{1}{2}q_3 - \tfrac{1}{2}q_4.$$

We shall show that q_4 is equal to $\mathbf{w}'[\mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21} \times \mathbf{I}]\mathbf{w}$. For each fixed value of $\mathbf{w}$ for which $\mathbf{A}$ exists there is an orthogonal $n \times n$ matrix $\mathbf{P}$ such that

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{I}_k & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{E}_k$$

by conditions (1) and (2). Partition $\mathbf{P}$ such that $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2]$ where $\mathbf{P}_1$ has dimension $n \times k$. Then $\mathbf{A} = \mathbf{P}_1\mathbf{P}_1'$ and $\mathbf{I} - \mathbf{A} = \mathbf{P}_2\mathbf{P}_2'$. Now

$$q_4 = \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{I})(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}$$
$$= \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{I})(\mathbf{I} \times \mathbf{P})(\mathbf{I} \times \mathbf{P})'(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})^{-1}$$
$$\cdot(\mathbf{I} \times \mathbf{P})(\mathbf{I} \times \mathbf{P})'(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}$$
$$= \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{I})(\mathbf{I} \times \mathbf{P})[(\mathbf{I} \times \mathbf{P})'(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{A})(\mathbf{I} \times \mathbf{P})]^{-1}$$
$$\cdot(\mathbf{I} \times \mathbf{P})'(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}$$

$$= \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{I})(\mathbf{I} \times \mathbf{P})(\mathbf{R}_{11} \times \mathbf{I} - 2i\mathbf{T} \times \mathbf{E}_k)^{-1}(\mathbf{I} \times \mathbf{P}')(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}$$

$$= \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{I})(\mathbf{I} \times \mathbf{P}) \begin{bmatrix} (\mathbf{R}_{11} - 2i\mathbf{T}) \times \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11} \times \mathbf{I}_{n-k} \end{bmatrix}^{-1}$$

$$\cdot (\mathbf{I} \times \mathbf{P}')(\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}$$

$$= \mathbf{w}'(\mathbf{R}_{21} \times \mathbf{P}) \begin{bmatrix} (\mathbf{R}_{11} - 2i\mathbf{T})^{-1} \times \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11}^{-1} \times \mathbf{I}_{n-k} \end{bmatrix} (\mathbf{R}_{12} \times \mathbf{P}')\mathbf{w}$$

$$= \mathbf{w}'(\mathbf{R}_{21} \times [\mathbf{P}_1, \mathbf{P}_2]) \begin{bmatrix} (\mathbf{R}_{11} - 2i\mathbf{T})^{-1} \times \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11}^{-1} \times \mathbf{I}_{n-k} \end{bmatrix} \left(\mathbf{R}_{12} \times \begin{bmatrix} \mathbf{P}_1' \\ \mathbf{P}_2' \end{bmatrix} \right) \mathbf{w}$$

$$= \mathbf{w}'[\mathbf{R}_{21} \times \mathbf{P}_1, \mathbf{R}_{21} \times \mathbf{P}_2] \begin{bmatrix} (\mathbf{R}_{11} - 2i\mathbf{T})^{-1} \times \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11}^{-1} \times \mathbf{I}_{n-k} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{12} \times \mathbf{P}_1' \\ \mathbf{R}_{12} \times \mathbf{P}_2' \end{bmatrix} \mathbf{w}$$

$$= \mathbf{w}'\{(\mathbf{R}_{21} \times \mathbf{P}_1)[(\mathbf{R}_{11} - 2i\mathbf{T})^{-1} \times \mathbf{I}_k](\mathbf{R}_{12} \times \mathbf{P}_1')$$

$$+ (\mathbf{R}_{21} \times \mathbf{P}_2)(\mathbf{R}_{11}^{-1} \times \mathbf{I}_{n-k})(\mathbf{R}_{12} \times \mathbf{P}_2')\}\mathbf{w}$$

$$= \mathbf{w}'[\mathbf{R}_{21}(\mathbf{R}_{11} - 2i\mathbf{T})^{-1}\mathbf{R}_{12} \times \mathbf{P}_1\mathbf{P}_1']\mathbf{w} + \mathbf{w}'[\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12} \times \mathbf{P}_2\mathbf{P}_2']\mathbf{w}$$

$$= \mathbf{w}'[\mathbf{R}_{21}(\mathbf{R}_{11} - 2i\mathbf{T})^{-1}\mathbf{R}_{12} \times \mathbf{A}]\mathbf{w} + \mathbf{w}'[\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12} \times (\mathbf{I} - \mathbf{A})]\mathbf{w}.$$

Consider the quantity $\mathbf{w}'(\mathbf{B} \times \mathbf{A})\mathbf{w}$ where $\mathbf{B} = \mathbf{R}_{21}(\mathbf{R}_{11} - 2i\mathbf{T})^{-1}\mathbf{R}_{12}$. We get

$$\begin{aligned} \mathbf{w}'(\mathbf{B} \times \mathbf{A})\mathbf{w} &= \operatorname{tr}[\mathbf{w}'(\mathbf{B} \times \mathbf{A})\mathbf{w}] = \operatorname{tr}[\textstyle\sum_i \sum_j (\mathbf{w}_i'\mathbf{B}a_{ij}\mathbf{w}_j)] \\ &= \textstyle\sum\sum \operatorname{tr}(\mathbf{w}_i'\mathbf{B}a_{ij}\mathbf{w}_j) = \sum\sum \operatorname{tr}(B\mathbf{w}_j a_{ij}\mathbf{w}_i') \\ &= \operatorname{tr}[\mathbf{B} \textstyle\sum\sum \mathbf{w}_j a_{ij}\mathbf{w}_i'] \\ &= \operatorname{tr}[\mathbf{BWAW}'] = \operatorname{tr}[\mathbf{R}_{21}(\mathbf{R}_{11} - 2i\mathbf{T})^{-1}\mathbf{R}_{12}\mathbf{WAW}'] = 0 \end{aligned}$$

by condition (3) where $\mathbf{W} = [\mathbf{w}_i, \cdots, \mathbf{w}_n]$. Hence $q_4 = \mathbf{w}'(\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12} \times \mathbf{I})\mathbf{w}$ and the integral in (3.3) becomes

$$\int_{\infty}^{\infty} \cdots \int_{\infty}^{\infty} \exp\{-\tfrac{1}{2}\mathbf{w}'[(\mathbf{R}_{22} - \mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12}) \times \mathbf{I}]\mathbf{w}\}\, d\mathbf{w}$$

$$= (2\pi)^{np_2/2}|\mathbf{R}_{22} - \mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12}|^{-n/2}$$

and the characteristic function is $\phi_S(\mathbf{t}) = |\mathbf{I} - 2i\mathbf{R}_{11}^{-1}\mathbf{T}|^{-k/2}$ and this is the characteristic function of $W(p_1, k, \mathbf{V}_{11.2})$ so the theorem is proved.

4. Illustrations.

EXAMPLE 4.1. Consider the general linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ where $\mathbf{y}$ is an $n \times 1$ random vector, $\mathbf{X}$ is an $n \times p$ matrix of constants of rank $k \leqq p < n$, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters and $\mathbf{e}$ is an $n \times 1$ unobservable random normal vector with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}$. Define $\mathbf{Q}'$ to be a $t \times n$ matrix whose elements are functions of the p elements in $\mathbf{X}'\mathbf{y}$ and suppose $\mathbf{Q}$ is such that each element exists with probability one and the rank of $\mathbf{Q}'[\mathbf{I}-$

$\mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}']\mathbf{Q}$ is equal to the positive integer m with probability one. Define $\mathbf{A}$ and $\mathbf{B}$ by

$$\mathbf{A} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}']\mathbf{Q}[\mathbf{Q}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}')\mathbf{Q}]^c\mathbf{Q}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}']$$
$$\mathbf{B} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}'] - \mathbf{A}$$

where $\mathbf{F}^c$ is a conditional inverse of $\mathbf{F}$, i.e., $\mathbf{F}\mathbf{F}^c\mathbf{F} = \mathbf{F}$. We identify matrices $\mathbf{K}$ and $\mathbf{L}$ as: $\mathbf{K} = \mathbf{X}'$; $\mathbf{L} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}'$. Clearly $\mathbf{L}\mathbf{K}' = \mathbf{0}$. Therefore the elements of $\mathbf{A}$ and $\mathbf{B}$ are functions of $\mathbf{K}\mathbf{y}$ and the following are easily verified (where appropriate, the conditions hold "with probability one").

$$\begin{aligned} \mathbf{L}\mathbf{A}\mathbf{L}' &= \mathbf{A} & \mathbf{A}\mathbf{B} &= \mathbf{0} \\ \mathbf{L}\mathbf{B}\mathbf{L}' &= \mathbf{B} & \text{rank }(\mathbf{A}) &= m \\ \mathbf{A} &= \mathbf{A}^2 & \text{rank }(\mathbf{B}) &= n - k - m \\ \mathbf{B} &= \mathbf{B}^2 & [\mathcal{E}(\mathbf{y})]'\mathbf{A}[\mathcal{E}(\mathbf{y})] &= 0 \\ & & [\mathcal{E}(\mathbf{y})]'\mathbf{B}[\mathcal{E}(\mathbf{y})] &= 0 \end{aligned}$$

By Theorems 3.1 and 3.2 it follows that $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are distributed as independent central chi-square random variables with m and $n - k - m$ degrees of freedom respectively.

EXAMPLE 4.2. Consider the two-way classification model

$$y_{ij} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + e_{ij}; \qquad i = 1, 2, \cdots, t; \quad j = 1, 2, \cdots, b,$$

where $\sum_i \tau_i = \sum_j \gamma_j = \sum_i (\tau\gamma)_{ij} = \sum_j (\tau\gamma)_{ij} = 0$ and where the e_{ij} are jointly independent and $e_{ij} \sim N(0, \sigma^2)$. Tukey [6] devised a test of the hypothesis H_0: $(\tau\gamma)_{ij} = 0$ for all i and j. The numerator of the test statistic is s_1^2 where

$$s_1^2 = \frac{[\sum_i \sum_j (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})(y_{i\cdot} - y_{\cdot\cdot})(y_{\cdot j} - y_{\cdot\cdot})]^2}{\sum_i (y_{i\cdot} - y_{\cdot\cdot})^2 \sum_j (y_{\cdot j} - y_{\cdot\cdot})^2}.$$

It is straightforward to show that s_1^2 is a special case of $\mathbf{y}'\mathbf{A}\mathbf{y}$ in Example 4.1. The denominator of the test devised by Tukey is $s_2^2 - s_1^2$ where

$$s_2^2 = \sum_i \sum_j (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2$$

and it is straightforward to show that $s_2^2 - s_1^2$ is a special case of $\mathbf{y}'\mathbf{B}\mathbf{y}$ in Example 4.1. From these facts we can obtain the distribution of $s_1^2/(s_2^2 - s_1^2)$ as a central F distribution under H_0.

EXAMPLE 4.3. Let $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n$ be independent $p \times 1$ vectors where $\mathbf{y}_i \sim N(\mathbf{0}, \mathbf{V})$ where $\mathbf{V}$ is a $p \times p$ positive definite matrix with $p < n$. We define $\mathbf{S}$ by $\mathbf{S} = \sum_{i=1}^n \mathbf{y}_i\mathbf{y}_i'$ and $\mathbf{S} \sim W(n, p, \mathbf{V})$. If we *partition* $\mathbf{S}$ and $\mathbf{V}$ by

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}; \qquad \mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$$

where $\mathbf{V}_{11}$ and $\mathbf{S}_{11}$ are $p_1 \times p_1$ matrices then $\mathbf{S}_{11} - \mathbf{S}_{12}\mathbf{S}_{22}^{-1}\mathbf{S}_{21} \sim W(p_1, n - p_2, \mathbf{V}_{11.2})$ where $\mathbf{V}_{11.2} = \mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$.

This result is verified by defining $\mathbf{u}_i$ and $\mathbf{w}_i$ by $\mathbf{y}_i' = [\mathbf{u}_i', \mathbf{w}_i']$ and noticing that $\mathbf{S}_{11} = \sum_{i=1}^n \mathbf{u}_i\mathbf{u}_i'$; $\mathbf{S}_{12} = \sum_{j=1}^n \mathbf{u}_j\mathbf{w}_j'$; $\mathbf{S}_{22} = \sum_{t=1}^n \mathbf{w}_t\mathbf{w}_t'$ and therefore

$$\begin{aligned}\mathbf{S}_{11} - \mathbf{S}_{12}\mathbf{S}_{22}^{-1}\mathbf{S}_{21} &= \textstyle\sum_i \mathbf{u}_i\mathbf{u}_i' - (\sum_i \mathbf{u}_i\mathbf{w}_i')(\sum_t \mathbf{w}_t\mathbf{w}_t')^{-1}(\sum_j \mathbf{w}_j\mathbf{u}_j')\\ &= \textstyle\sum_i\sum_j \mathbf{u}_i[\delta_{ij} - \mathbf{w}_i' \ (\sum_t \mathbf{w}_t\mathbf{w}_t')^{-1}\mathbf{w}_j]\mathbf{u}_j'\\ &= \textstyle\sum_i\sum_j \mathbf{u}_i\mathbf{u}_j' a_{ij},\end{aligned}$$

where $\mathbf{A} = [a_{ij}]$ is a function of $\mathbf{w}_t$; $t = 1, 2, \cdots, n$. If we define $\mathbf{W}$ by $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_n]$, then $\mathbf{A} = \mathbf{I} - \mathbf{W}'(\mathbf{W}\mathbf{W}')^{-1}\mathbf{W}$ and the following clearly hold with probability one:

$$\mathbf{A} = \mathbf{A}^2; \qquad \operatorname{tr}(\mathbf{A}) = n - p_2; \qquad [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n]\mathbf{A} = \mathbf{0}.$$

Hence we use Theorem 3.3 and the result follows.

EXAMPLE 4.4. Let the 3×1 random vector $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I})$ and consider $\mathbf{y}'\mathbf{A}\mathbf{y}$ where $\mathbf{A}$ is defined by

$$\mathbf{A} = c^{-1} \cdot \begin{bmatrix} \frac{1}{2}(y_1 - y_2)^2 & \frac{1}{2}(y_1 - y_2)^2 & 2^{-\frac{1}{2}}(y_1 - y_2)\log|y_1 - y_2| \\ \frac{1}{2}(y_1 - y_2)^2 & \frac{1}{2}(y_1 - y_2)^2 & 2^{-\frac{1}{2}}(y_1 - y_2)\log|y_1 - y_2| \\ 2^{-\frac{1}{2}}(y_1 - y_2)\log|y_1 - y_2| & 2^{-\frac{1}{2}}(y_1 - y_2)\log|y_1 - y_2| & \log^2|y_1 - y_2| \end{bmatrix}$$

where $c = [(y_1 - y_2)^2 + \log^2|y_1 - y_2|]$. If we define $\mathbf{L}$, and $\mathbf{K}$ by

$$\mathbf{L} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}; \qquad \mathbf{K} = [1 \ -1 \ 0]$$

then $\mathbf{L}\mathbf{K}' = \mathbf{0}$ and the following hold with probability one.

(1) $\mathbf{L}'\mathbf{A}\mathbf{L} = \mathbf{A}$;
(2) $\mathbf{A} = \mathbf{A}^2$;
(3) $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$;
(4) rank $(\mathbf{A}) = 1$.

Hence by Theorem 3.1 $\mathbf{y}'\mathbf{A}\mathbf{y}$ is distributed as chi-square with one degree of freedom.

Examples 4.2 and 4.3 are well known results that are easily proved by using Theorems 3.1, 3.2 and 3.3 of this paper. Example 4.1 will be used to prove a number of new and useful results in another paper.

Many of the theorems that concern quadratic forms of normal variables that are useful in the theory of the general linear model can be extended to include cases when the matrices of the quadratic forms have random elements. This will be the topic of a further paper.

REFERENCES

[1] GRAYBILL, FRANKLIN A. (1961). *An Introduction to Linear Statistical Models*, **1**. McGraw-Hill, New York.

[2] GRAYBILL, F. A. and MARSAGLIA, G. (1957). Idempotent matrices and quadratic forms in the general linear hypothesis. *Ann. Math. Statist.* **28** 678–686.

[3] HOGG, R. V. and CRAIG, A. T. (1959). On the decomposition of certain χ^2 variables. *Ann. Math. Statist.* **29** 608–610.

[4] LUTHER, NORMAN Y. (1965). Decomposition of symmetric matrices and distributions of quadratic forms. *Ann. Math. Statist.* **36** 683–690.

[5] RAO, C. RADHAKRISHNA. (1965). *Linear Statistical Inference and Its Applications*. Wiley, New York.

[6] TUKEY, J. W. (1949). One degree of freedom for nonadditivity. *Biometrics* **5** 232–242.