QUADRATIC FORMS IN NORMAL VARIABLES

By

Issie Scarowsky

Department of Mathematics
McGill University
Montreal.

March 1973

QUADRATIC FORMS IN NORMAL VARIABLES

by

Issie Scarowsky

Abstract

This thesis gives a comprehensive account of results in quadratic forms in normal variables. The characteristic function of an arbitrary quadratic form in normal variables, necessary and sufficient conditions for a quadratic form in normal variables to have a chi-square distribution, and necessary and sufficient conditions for two quadratic forms in normal variables to be independent are obtained. These results are extended to include nonhomogeneous quadratic forms and bilinear forms in normal variables. The historical evolvement of these results is traced and an exhaustive bibliography is included.

Department of Mathematics,
McGill University,
Montreal.

M.Sc.

March 1973

QUADRATIC FORMS IN NORMAL VARIABLES

par

Issie Scarowsky

Résumé

Cette thèse est un compte rendu approfondi de résultats sur les formes bilinéaires pour variables aléatoires gaussiennes. Ainsi, nous obtenons la fonction caractéristique d'une forme bilinéaire quelconque, des conditions nécessaires et suffisantes pour qu'une forme bilinéaire suive une loi chi-carré, et des conditions nécessaires et suffisantes pour l'indépendance de deux formes bilinéaires. Ces résultats sont généralisés pour des formes bilinéaires non-homogènes. Nous retraçons aussi le développement historique de ces résultats et nous incluons une bibliographie détaillée.

Département de Mathématiques,
Université McGill,
Montréal.

M.Sc.

Mars 1973

QUADRATIC FORMS IN NORMAL VARIABLES

by

Issie Scarowsky

A thesis submitted to the Faculty of Graduate Studies and Research, in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics,
McGill University,
Montreal.

March 1973

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to George P.H. Styan for suggesting the topic of my thesis and for providing invaluable suggestions and guidance throughout. He was also most helpful in the compilation of the bibliography.

I would also like to thank Leslie Ann Hathaway for her continuous encouragement and for the many aspects in the preparation of this thesis which she did so well. Thanks are due to Lesley Walker for helping with the bibliography.

I would like to thank my fellow students and colleagues at McGill and my co-workers at Bell Canada and, in particular, Dave Hartman for their continuous moral support and understanding during the writing of this thesis.

Finally, this thesis was typed by Heather Benson who did a superb job for which I am most grateful.

TABLE OF CONTENTS

		Page
Chapter I.	Introduction and Preliminaries	1
Chapter II.	Characteristic Function, Cumulant Generating Function; Cumulants and Moments	5
Chapter III.	Chi-Squaredness	21
Chapter IV.	Independence	40
Chapter V.	Cochran's Theorem	63
References and Bibliography	....	75

CHAPTER I

Introduction and Preliminaries

There have been many papers and articles throughout the statistical literature on quadratic forms in normal variables. This literature appears widely with much duplication of results. The purpose of this thesis is to organize these results, to obtain concise and complete proofs of the basic theorems, and in a historical perspective to trace their evolution. In addition, many extensions are given and some new results are also presented.

This thesis examines, in detail, the following topics on quadratic forms in normal variables. The first area covered concerns the characteristic function, the cumulants and the moments of an arbitrary quadratic form in normal variables. The second topic is the derivation of necessary and sufficient conditions for a quadratic form in normal variables to follow a chi-square distribution. The third subject concerns the necessary and sufficient conditions for two quadratic forms in normal variables to be independent. This result is often referred to as Craig's Theorem. The last subject discussed is the interaction between quadratic forms which individually follow chi-square distributions and their independence. This is often referred to as Cochran's Theorem.

These results are extended to include nonhomogeneous quadratic forms and bilinear forms in normal variables. In addition, we have considered what these results reduce to when there are additional conditions on the matrix of the quadratic form or on the distribution of the normal variables. Also included is a bibliography which contains over a hundred references covering exhaustively the above four subjects.

One area this thesis does not cover is the computation of the distribution function (or approximations thereof) of a quadratic form in normal variables. For references on this and other aspects of quadratic forms in normal variables see *A Bibliography of Multivariate Statistical Analysis* by Anderson, Das Gupta, and Styan (1972), under subject-matter code 2.5.

We shall now discuss the notation used in this thesis as well as some matrix results that have been assumed. Both Mirsky (1955) and Searle (1966) are good references for these results.

Throughout this thesis we use matrix notation. Capital letters denote matrices and underscored lower case letters denote column vectors. Transposition is denoted by a prime, with row vectors always appearing primed. For any square matrix A, we use rk(A) to denote its rank, tr(A) its trace, $|A|$ its determinant, and $ch_i(A)$ its *ith* largest characteristic root.

A quadratic form in normal variables is a quadratic expression in random variables which follow a multivariate normal distribution. Suppose the random vector $\underset{\sim}{x} = (x_1,\ldots,x_p)'$ follows a multivariate normal distribution. Then $Q = \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij}x_ix_j$ is a quadratic form in normal variables. If we let the matrix $A = \{a_{ij}\}$, then Q may be written $Q = \underset{\sim}{x}'A\underset{\sim}{x}$. The matrix of the quadratic form may always be assumed to be symmetric for $Q = \underset{\sim}{x}'A\underset{\sim}{x} = \underset{\sim}{x}'A'\underset{\sim}{x} = \underset{\sim}{x}' \frac{(A+A')}{2} \underset{\sim}{x}$, and $\frac{(A+A')}{2}$ is symmetric. A nonhomogeneous quadratic form contains quadratic, linear and constant terms and may be expressed as

$\underset{\sim}{x}'A\underset{\sim}{x} + \underset{\sim}{b}'\underset{\sim}{x} + c$. A bilinear form is a quadratic expression involving a sum of crossproducts between two distinct sets of variables. For example, if $\underset{\sim}{x} = (x_1,x_2,\ldots,x_p)'$ and $\underset{\sim}{y} = (y_1,y_2,\ldots,y_q)'$ then $Q = \sum_{i=1}^{p}\sum_{j=1}^{q} a_{ij}x_iy_j = \underset{\sim}{x}'A\underset{\sim}{y}$, with A $p\times q$, is a bilinear form. Throughout this thesis we assume that all matrices are real.

If the $p\times p$ matrix A has $rk(A) = p$ then A is nonsingular and A^{-1} exists; otherwise A is singular. If A is symmetric, then its characteristic roots are all real. Moreover, if λ_i is a characteristic root of A and if $P(A)$ is a polynomial expression in A then $P(\lambda_i)$ is a characteristic root of $P(A)$. Also the trace $tr(A) = \sum_{i=1}^{p} ch_i(A)$ and the determinant $|A| = \prod_{i=1}^{p} ch_i(A)$. If A and B are two matrices such that both AB and BA are defined, then the nonzero $ch(AB)$ equal the nonzero $ch(BA)$. As a result $tr(AB) = tr(BA)$ and $|I-AB|=|I-BA|$, where I is the identity matrix (of appropriate order). The rank of a square matrix A is greater than or equal to the number of nonzero $ch(A)$; if A is symmetric, however, then $rk(A)$ equals the number of nonzero $ch(A)$.

A symmetric matrix A is positive definite if $\underset{\sim}{x}'A\underset{\sim}{x} > 0$ for all $\underset{\sim}{x} \neq \underset{\sim}{0}$; the matrix A is positive semidefinite if $\underset{\sim}{x}'A\underset{\sim}{x} \geq 0$ for all $\underset{\sim}{x}$. The symmetric matrix A is positive semidefinite if and only if all $ch(A) \geq 0$ and is positive definite if and only if all $ch(A) > 0$. Every positive definite matrix is nonsingular. For any $p\times p$ symmetric matrix A with $rk(A) = r$ there exists an orthogonal matrix P such that $A = P\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}P'$ where Λ is an $r\times r$ diagonal matrix containing the nonzero characteristic roots of A. If A is positive semidefinite, then there exists a real matrix T of full column rank such that $A = TT'$. For any matrix B, the trace $tr(BB')$ equals the sum of squares of the elements in B. Thus $tr(BB') = 0$ implies that the matrix $B=0$.

A matrix A is said to be idempotent if $A^2=A$. Throughout this thesis, idempotent matrices are assumed to be symmetric though the following results also hold without symmetry. The characteristic roots of an idempotent matrix are 1 or 0. In fact, if $r = rk(A)$ there are exactly r characteristic roots equal to 1, with the remaining roots all 0. Thus for an idempotent matrix $rk(A) = tr(A)$. Also, symmetric idempotent matrices are positive semidefinite.

CHAPTER II

Characteristic Function, Cumulant Generating Function; Cumulants and Moments.

Characteristic functions, and moment generating functions when they exist, are a powerful tool in probability theory and in mathematical statistics. They may be used in identifying the distribution of a random variable and in determining its moments. In addition, independence of random variables may be established by characteristic functions.

Wilks (1962), in an introductory paragraph to the chapter on characteristic functions, states

> "One of the most important classes of problems in mathematical statistics is the determination of distribution functions of measurable functions of random variables... . Some situations, particularly those involving linear functions of independent random variables, can often be handled in an elegant manner by making use of the characteristic function of the particular function of the random variable under consideration. The characteristic function [is] also useful for such tasks as generating moments and cumulants of distributions and testing independence of two or more functions of random variables."

In the context of quadratic forms in normal variables, the merits of these techniques are brought out rather well. In particular, the probability density function and the cumulative distribution function of an arbitrary quadratic function in normal variables are not available in closed form, although approximations have been found. [See, e.g., Imhof (1961).] The characteristic function, however, is known in closed form and it is effectively the only way in which we may prove results concerning chi-squaredness and independence.

In the next chapter, necessary and sufficient conditions are obtained with the use of characteristic functions for a quadratic form in normal variables to follow a chi-square distribution. The characteristic function is also used in establishing conditions for a sum of quadratic forms to follow a chi-square distribution. Also, necessary and sufficient conditions for quadratic forms to be independent are found by equating the joint characteristic function to the product of the marginals. For these reasons, we shall first determine the characteristic function of an arbitrary quadratic form in normal variables.

Let $\underset{\sim}{x}$ be a $p\times 1$ random column vector distributed normally with mean vector $\underset{\sim}{\mu}$ and variance-covariance matrix $\underset{\sim}{\Sigma}$, possibly singular. In what follows this will be denoted $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ and where the dimension is understood to be p the subscript may be omitted. Then $x_j \sim N_1(\mu_j,\sigma_{jj})$, $j = 1,\ldots,p$, where $\underset{\sim}{x} = \{x_j\}$, $\underset{\sim}{\mu} = \{\mu_j\}$, $\underset{\sim}{\Sigma} = \{\sigma_{ij}\}$.

The following results are standard and stated without proof:

$$f(x_j) = (2\pi\sigma_{jj})^{-1/2}\, e^{-(x_j-\mu_j)^2/(2\sigma_{jj})} \tag{2.1}$$

$$\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\underset{\sim}{\Sigma}) \Leftrightarrow \underset{\sim}{\ell}'\underset{\sim}{x} \sim N_1(\underset{\sim}{\ell}'\underset{\sim}{\mu},\underset{\sim}{\ell}'\underset{\sim}{\Sigma}\underset{\sim}{\ell}) \quad \forall\ \underset{\sim}{\ell}\ p\times 1 \tag{2.1a}$$

$$g(\underset{\sim}{x}) = (2\pi)^{-p/2}\, |\underset{\sim}{\Sigma}|^{-1/2}\, e^{-(\underset{\sim}{x}-\underset{\sim}{\mu})'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})/2} \quad ; \ |\underset{\sim}{\Sigma}| > 0 \tag{2.2}$$

$$E(e^{itx_j}) = e^{it\mu_j - t^2\sigma_{jj}/2} \tag{2.3}$$

$$E(e^{i\mathbf{t}'\mathbf{x}}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2} . \tag{2.4}$$

In (2.1) and (2.2) $f(\cdot)$ and $g(\cdot)$ denote density functions; in (2.2) $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$ while in (2.3) and (2.4) $E(\cdot)$ denotes mathematical expectation.

We now present the characteristic function of a symmetric quadratic form in normal variables. We do not exclude the possibility of a singular variance-covariance matrix. This will allow us to obtain the characteristic function of a nonhomogeneous quadratic form and of a bilinear form by simple transformations to homogeneous quadratic forms in a singular normal vector. See, for example, Corollaries 2.1 and 2.2 .

THEOREM 2.1: *If* A *is a real symmetric* $p\times p$ *matrix and* $\mathbf{x} \sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ *positive semidefinite, then the characteristic function of* $\mathbf{x}'A\mathbf{x}$ *is*

$$E(e^{it\mathbf{x}'A\mathbf{x}}) = \frac{e^{it\boldsymbol{\mu}'(I-2itA\boldsymbol{\Sigma})^{-1}A\boldsymbol{\mu}}}{|I-2itA\boldsymbol{\Sigma}|^{1/2}} . \tag{2.5}$$

We shall prove this theorem by obtaining the moment generating function of $\mathbf{x}'A\mathbf{x}$, i.e., $E(e^{t\mathbf{x}'A\mathbf{x}})$. We then use the general result [cf. e.g., Lukacs (1970), p. 11 and Section 7.1] that if the moment generating function, $M(t)$, of a random variable exists in a strip about the origin, then the characteristic function of that random variable is given by $f(t) = M(it)$ for all real t.

In order to simplify the proof of the above and other results we use the following lemmas.

LEMMA 2.1: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *then for any real* $m \times p$ *matrix* H, $H\underset{\sim}{x} \sim N_m(H\underset{\sim}{\mu}, H\Sigma H')$.

Proof: Immediate from (2.1a) and (2.4).

LEMMA 2.2: *If* $x \sim N_1(0,1)$, *then*

$$E[e^{sx^2 + tx}] = \frac{e^{t^2/[2(1-2s)]}}{(1-2s)^{1/2}} \quad ; \quad s < 1/2 . \tag{2.6}$$

Proof: Since $x \sim N_1(0,1)$, we have

$$E[e^{sx^2 + tx}] = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{sx^2 + tx - x^2/2} dx$$

$$= (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-[x^2(1-2s) - 2xt]/2} dx. \tag{2.7}$$

Substitute $u = x(1-2s)^{1/2} - t(1-2s)^{-1/2}$ which is real when $s < 1/2$. Then $u^2 = x^2(1-2s) + t^2(1-2s)^{-1} - 2xt$ so that

$$E[e^{sx^2 + tx}] = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-u^2/2} du (1-2s)^{-1/2} e^{t^2/[2(1-2s)]}$$

$$= \frac{e^{t^2/[2(1-2s)]}}{(1-2s)^{1/2}} .$$

We notice that when $s = 0$, (2.6) gives the moment generating function of $x \sim N_1(0,1)$,

$$E(e^{tx}) = e^{t^2/2} .$$

When $t = 0$, (2.6) gives the moment generating function of a χ_1^2,

$$E[e^{sx^2}] = E[e^{s\chi_1^2}] = (1-2s)^{-1/2} \; ; \quad s < 1/2 . \tag{2.8}$$

The moment generating function of a noncentral $\chi_1^2(\mu^2)$ is also readily obtained. Let $w = x + \mu$, where $x \sim N_1(0,1)$. Then $w^2 = x^2 + 2\mu x + \mu^2 \sim \chi_1^2(\mu^2)$ and

$$
\begin{aligned}
E[e^{sw^2}] &= e^{s\mu^2} E[e^{sx^2 + 2\mu sx}] \\
&= \frac{e^{s\mu^2} e^{2s^2\mu^2/(1-2s)}}{(1-2s)^{1/2}} \\
&= \frac{e^{s\mu^2/(1-2s)}}{(1-2s)^{1/2}} \; ; \quad s < \frac{1}{2} \, . \qquad (2.9)
\end{aligned}
$$

LEMMA 2.3: *If* B *and* C *are two* $p \times r$ *matrices such that no characteristic roots of* BC' *are* 1, *then*

$$(I-BC')^{-1} = I + B(I-C'B)^{-1}C' \, . \qquad (2.10)$$

<u>Proof</u>: Since BC' has no characteristics roots of 1 the inverses in (2.10) exist. The product

$$
\begin{aligned}
&[I + B(I-C'B)^{-1}C'] \, [I - BC'] \\
&\quad = I - BC' + B(I-C'B)^{-1}C' - B(I-C'B)^{-1}C'BC' \\
&\quad = I - BC' + B(I-C'B)^{-1} \, (I-C'B)C' \\
&\quad = I
\end{aligned}
$$

and (2.10) is proved.

LEMMA 2.4: [Anderson (1958), p. 25] *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *with* $rk(\Sigma) = r \leq p$, *then there exists a random vector* $\underset{\sim}{y} \sim N_r(\underset{\sim}{0}, I_r)$ *and a* *real* p×r *matrix* A *such that* $\Sigma = AA'$ and

$$x = Ay + \mu \; . \tag{2.11}$$

Proof: There exists a real p×p nonsingular matrix T such that $T\Sigma T' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. For example, we can write $\Sigma = P \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix} P'$, where Λ_r is a diagonal matrix containing the r nonzero characteristic roots of Σ (which are all positive). Then

$$T = P \begin{pmatrix} \Lambda_r^{-1/2} & 0 \\ 0 & I_{p-r} \end{pmatrix} P' \tag{2.12}$$

is such a matrix. Let $w = Tx$, then by Lemma 2.1

$$w \sim N_p(T\mu, \; T\Sigma T') \; . \tag{2.13}$$

Write $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, where w_1 is r×1 and w_2 is (p-r)×1 and write $v = T\mu = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, where v_1 is r×1 and v_2 is (p-r)×1. Since $T\Sigma T' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, we have $w_1 \sim N_r(v_1, I_r)$ and $w_2 \equiv v_2$ with probability 1. Let $T^{-1} = [A,B]$, where A is p×r and B is p×(p-r). Then $x = T^{-1}w$ and so

$$\begin{aligned} x &= Aw_1 + Bw_2 = Aw_1 + Bv_2 \\ &= A(w_1 - v_1) + Av_1 + Bv_2 \\ &= A(w_1 - v_1) + T^{-1}v \\ &= Ay + \mu \; , \end{aligned} \tag{2.14}$$

where $\underset{\sim}{y} = \underset{\sim}{w}_1 - \underset{\sim}{v}_1$. Clearly $\underset{\sim}{y} \sim N_r(\underset{\sim}{0}, I)$. Also, since $T\Sigma T' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$,

$$\Sigma = T^{-1} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} (T^{-1})'$$

$$= [A,B] \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A' \\ B' \end{pmatrix}$$

$$= AA' , \tag{2.15}$$

and the lemma is proved.

Proof of Theorem 2.1: Using Lemma 2.4, there exists a $\underset{\sim}{y}$ such that $\underset{\sim}{x} = T\underset{\sim}{y} + \underset{\sim}{\mu}$, where $\underset{\sim}{y} \sim N_r(\underset{\sim}{0}, I)$ and $TT' = \Sigma$. Also, there exists an orthogonal matrix P such that $T'AT = P'\Lambda P$; Λ is an $r \times r$ diagonal matrix containing the characteristic roots, $\lambda_1, \lambda_2, \ldots, \lambda_r$, of $T'AT$, or of $A\Sigma$ (except for an additional $p-r$ zero roots).

Let $\underset{\sim}{z} = P\underset{\sim}{y}$ and denote the *ith* component of $\underset{\sim}{z}$ by z_i; by Lemma 2.1 $\underset{\sim}{z} \sim N_r(\underset{\sim}{0}, I)$. Also, let $\underset{\sim}{\nu}' = 2\underset{\sim}{\mu}'ATP'$; then

$$\underset{\sim}{x}'A\underset{\sim}{x} = \underset{\sim}{y}'T'AT\underset{\sim}{y} + 2\underset{\sim}{\mu}'AT\underset{\sim}{y} + \underset{\sim}{\mu}'A\underset{\sim}{\mu}$$

$$= \underset{\sim}{y}'P'\Lambda P\underset{\sim}{y} + 2\underset{\sim}{\mu}'AT\underset{\sim}{y} + \underset{\sim}{\mu}'A\underset{\sim}{\mu}$$

$$= \underset{\sim}{z}'\Lambda\underset{\sim}{z} + \underset{\sim}{\nu}'\underset{\sim}{z} + \underset{\sim}{\mu}'A\underset{\sim}{\mu}$$

$$= \sum_{i=1}^{r} (\lambda_i z_i^2 + \nu_i z_i) + \underset{\sim}{\mu}'A\underset{\sim}{\mu} .$$

The moment generating function of $\underset{\sim}{x}'A\underset{\sim}{x}$ is

$$E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}}) = E(e^{t\sum_{i=1}^{r}(\lambda_i z_i^2 + \nu_i z_i)})\, e^{t\underset{\sim}{\mu}'A\underset{\sim}{\mu}}$$

$$= E(\prod_{i=1}^{r} e^{t(\lambda_i z_i^2 + \nu_i z_i)})\, e^{t\underset{\sim}{\mu}'A\underset{\sim}{\mu}} . \qquad (2.16)$$

But since the z_i are independent, (2.16) becomes

$$E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}}) = [\prod_{i=1}^{r} E(e^{t(\lambda_i z_i^2 + \nu_i z_i)})]\, e^{t\underset{\sim}{\mu}'A\underset{\sim}{\mu}}$$

$$= [\prod_{i=1}^{r} \frac{e^{\frac{1}{2}\left[\frac{t^2\nu_i^2}{1-2t\lambda_i}\right]}}{(1-2t\lambda_i)^{1/2}}]\, e^{t\underset{\sim}{\mu}'A\underset{\sim}{\mu}} , \qquad (2.17)$$

using Lemma 2.2, provided $t\lambda_i < 1/2$ for $i = 1,2,\ldots,r$. This condition is satisfied for all the zero roots; it will be satisfied for all nonzero λ_i for all t such that $-k < t < k$ where k is the minimum value of $|(2\lambda_i)^{-1}|$ for all $i = 1,2,\ldots,r$ such that $\lambda_i \neq 0$.

We have

$$\prod_{i=1}^{r}(1-2t\lambda_i)^{1/2} = \prod_{i=1}^{p}[1-2tch_i(A\not{Z})]^{1/2}$$

$$= \prod_{i=1}^{p}[ch_i(I-2tA\not{Z})]^{1/2} = |I-2tA\not{Z}|^{1/2} , \qquad (2.18)$$

where $ch_i(\cdot)$ denotes the *ith* largest nonzero characteristic root, and also

$$\prod_{i=1}^{r} e^{\frac{1}{2}\left[\frac{t^2\nu_i^2}{1-2t\lambda_i}\right]} = e^{\frac{1}{2}t^2 \sum_{i=1}^{r} \frac{\nu_i^2}{1-2t\lambda_i}}$$

$$= e^{\frac{1}{2}t^2 \underset{\sim}{\nu}'(I-2t\Lambda)^{-1}\underset{\sim}{\nu}} . \tag{2.19}$$

Using (2.18) and (2.19) in (2.17) yields

$$E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}}) = \frac{e^{t\underset{\sim}{\mu}'A\underset{\sim}{\mu} + \frac{1}{2}t^2\underset{\sim}{\nu}'(I-2t\Lambda)^{-1}\underset{\sim}{\nu}}}{|I-2tA\Sigma|^{1/2}} . \tag{2.20}$$

Substituting $2\underset{\sim}{\mu}'ATP'$ back for $\underset{\sim}{\nu}'$ and using $P'(I-2t\Lambda)^{-1}P = (I-2tP'\Lambda P)^{-1} = (I-2tT'AT)^{-1}$, we obtain

$$E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}}) = \frac{e^{t\underset{\sim}{\mu}'(I+2tAT(I-2tT'AT)^{-1}T')A\underset{\sim}{\mu}}}{|I-2tA\Sigma|^{1/2}} . \tag{2.21}$$

From Lemma 2.3, with $B = 2tAT$ and $C = T$, (2.21) yields $E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}}) = \dfrac{e^{t\underset{\sim}{\mu}'(I-2tA\Sigma)^{-1}A\underset{\sim}{\mu}}}{|I-2tA\Sigma|^{1/2}}$ for all real t such that $|t| < k$, where k is defined as in the paragraph just after (2.17). This implies (2.5) using the general results quoted above Lemma 2.1. Thus Theorem 2.1 is proven.

The above proof through (2.21) follows closely the derivation by Rohde, Urquhart, and Searle (1966).

Ogasawara and Takahashi (1951) give the moment generating function of $\underset{\sim}{x}'A\underset{\sim}{x}$, with $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$ (Σ positive semidefinite) as

$$E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}}) = |I-2tA_1|^{-1/2} e^{\underset{\sim}{\mu}'[At+2t^2A\Sigma^{1/2}(I-2tA_1)^{-1}\Sigma^{1/2}A]\underset{\sim}{\mu}} , \tag{2.21a}$$

where $\Sigma^{1/2}$ is the positive semidefinite square root of Σ and $A_1 = \Sigma^{1/2}A\Sigma^{1/2}$. This appears to be the first treatment where Σ could be singular though no proof of (2.21a) is given. Unfortunately, the journal in which Ogasawara and Takahashi published appears not to have been readily available so that (2.21a), as well as various results these authors derive from (2.21a), were later proved again.

The form (2.5) was obtained by Mäkeläinen (1966), using an argument based on stochastic convergence [the Cramér-Lévy continuity theorem, see e.g., Wilks (1962)] extending a result of Plackett (1960). Plackett derived the characteristic function of $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ positive definite, as

$$E(e^{it\underset{\sim}{x}'A\underset{\sim}{x}}) = |I-2tA\Sigma|^{-1/2}e^{-\frac{1}{2}\underset{\sim}{\mu}'\Sigma^{-1}\underset{\sim}{\mu} + \frac{1}{2}\underset{\sim}{\mu}'\Sigma^{-1}(\Sigma^{-1}-2itA)^{-1}\Sigma^{-1}\underset{\sim}{\mu}} ,$$

which simplifies directly to (2.5) using Lemma 2.3. The continuity theorem was previously employed by Good (1963), who gave the characteristic function of $\underset{\sim}{x}'A\underset{\sim}{x}$ for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\Sigma)$, Σ positive semidefinite. Good (1963) also gave the joint characteristic function for two non-homogeneous quadratic forms as well as for a finite number of quadratic expressions.

Less general results on the characteristic function of quadratic forms in normal variables were known much earlier. Cochran (1934) gave the characteristic function of $\underset{\sim}{x}'A\underset{\sim}{x}$ for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},I)$. Craig (1938) extended this result and obtained the joint characteristic function of s quadratic forms where $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\sigma^2 I)$. Aitken (1940) obtained the joint characteristic function of a linear and a quadratic form and also of two quadratic forms in $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\Sigma)$, Σ positive definite. This result was also obtained by Sakamoto (1944b) and Ogawa (1946b). Both Ogawa (1950) and Carpenter (1950) (apparently independently) obtained the characteristic function with $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},I)$.

Khatri (1961a) derived the joint moment generating function for two nonhomogeneous forms in nonsingular normal variables. The following year, Khatri (1962) obtained the moment generating function for a single nonhomogeneous quadratic form in singular normal variables.

As stated previously, the characteristic function of nonhomogeneous forms or bilinear forms or systems of quadratic forms may be obtained directly from the characteristic function of a single, possibly singular, quadratic form. We offer the following two corollaries as examples.

COROLLARY 2.1: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ *positive semidefinite, then the characteristic function of* $Q = \underset{\sim}{x}'A\underset{\sim}{x} + 2\underset{\sim}{b}'\underset{\sim}{x} + c$ *is*

$$E(e^{itQ}) = \frac{e^{it[(\underset{\sim}{\mu}+2it\Sigma\underset{\sim}{b})'(I-2itA\Sigma)^{-1}(A\underset{\sim}{\mu}+\underset{\sim}{b})+\underset{\sim}{b}'\underset{\sim}{\mu}+c]}}{|I-2itA\Sigma|^{1/2}} . \tag{2.22}$$

<u>Proof</u>: We construct a new random vector $\underset{\sim}{y} = \begin{pmatrix}\underset{\sim}{x}\\1\end{pmatrix} \sim N_{p+1}(\underset{\sim}{\mu}_o,\Sigma_o)$, where $\underset{\sim}{\mu}_o = \begin{pmatrix}\underset{\sim}{\mu}\\1\end{pmatrix}$ and $\Sigma_o = \begin{pmatrix}\Sigma & \underset{\sim}{0}\\ \underset{\sim}{0}' & 0\end{pmatrix}$, with $\underset{\sim}{0}$ a p×1 column vector of zeros. If $A_o = \begin{pmatrix}A & \underset{\sim}{b}\\ \underset{\sim}{b}' & c\end{pmatrix}$, then $Q = \underset{\sim}{y}'A_o\underset{\sim}{y}$. Thus

$$E(e^{itQ}) = E(e^{it\underset{\sim}{y}'A_o\underset{\sim}{y}}) = \frac{e^{it\underset{\sim}{\mu}_o'(I-2itA_o\Sigma_o)^{-1}A_o\underset{\sim}{\mu}_o}}{|I-2itA_o\Sigma_o|^{1/2}} . \tag{2.23}$$

Substituting for μ_o, A_o and Σ_o and simplifying (2.23) gives (2.22).

COROLLARY 2.2: *If* x, y *have a joint multinormal distribution such that* $x \sim N_p(\mu_x, \Sigma_{xx})$, $y \sim N_q(\mu_y, \Sigma_{yy})$, *where both* Σ_{xx} *and* Σ_{yy} *are positive semidefinite and the cross-covariance matrix of* x *and* y *is* $\Sigma_{xy} = \Sigma'_{yx}$, *then the characteristic function of* $x'A_oy$ *is given by* (2.5) *with*

$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \text{ and } \quad A = \frac{1}{2}\begin{pmatrix} 0 & A_o \\ A_o & 0 \end{pmatrix}.$$

Proof: The $(p+q)\times 1$ random vector $z = \begin{pmatrix} x \\ y \end{pmatrix} \sim N_{p+q}(\mu, \Sigma)$, and $z'Az = x'A_oy$. Use of (2.5) on $z'Az$ completes the proof.

We now obtain the cumulant generating function and cumulants of an arbitrary quadratic form in normal variables. The cumulant generating function is essentially the logarithm of the characteristic function. It is also known [Lukacs (1970), pp. 26-27] that when the moment generating function exists, the cumulant generating function equals the logarithm of the moment generating function. This is the approach that will be used here. The k*th* cumulant is then the coefficient of $t^k/k!$ in the power-series expansion of the cumulant generating function. There is a one-to-one correspondence between cumulants and moments; thus, if the cumulants up to order k are specified then so are the first k moments, and vice-versa. See, for example, Kendall and Stuart (1969), Vol. I, pp. 68-71.

We use the following two lemmas.

LEMMA 2.5: *If the characteristic roots of a square matrix are all less than* 1 *in absolute value then*

$$(I-G)^{-1} = I + G + G^2 + \dots \quad . \qquad (2.24)$$

<u>Proof</u>: See e.g., Mirsky (1955), p. 332.

LEMMA 2.6: *If the characteristic roots of a square* n×n *matrix* G *are all real and less than* 1 *in absolute value then*

$$\log(|I-G|) = -\mathrm{tr} \sum_{k=1}^{\infty} G^k/k \, . \qquad (2.25)$$

<u>Proof</u>: $\log(|I-G|) = \log[\prod_{i=1}^{n} (1-g_i)] = \sum_{i=1}^{n} \log(1-g_i) = \sum_{i=1}^{n} (-\sum_{k=1}^{\infty} g_i^k/k)$ $= -\sum_{k=1}^{\infty} \sum_{i=1}^{n} g_i^k/k = -\sum_{k=1}^{\infty} \mathrm{tr}(G^k)/k$ where g_i, $i=1,\dots,n$ are the characteristic roots of G.

THEOREM 2.2: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ *positive semidefinite, and* A *is a real symmetric matrix, then the cumulant generating function of* $\underset{\sim}{x}'A\underset{\sim}{x}$ *is*

$$\phi(t) = \sum_{j=1}^{\infty} \frac{t^j}{j!} (j!2^{j-1}[\underset{\sim}{\mu}'(A\Sigma)^{j-1}A\underset{\sim}{\mu} + \frac{\mathrm{tr}(A\Sigma)^j}{j}]) \, . \qquad (2.26)$$

<u>Proof</u>: Using (2.5), $\phi(t) = \log[E(e^{t\underset{\sim}{x}'A\underset{\sim}{x}})] = t\underset{\sim}{\mu}'(I-tA\Sigma)^{-1}A\underset{\sim}{\mu}-\frac{1}{2}\log|I-tA\Sigma|$. We can find a $t_o > 0$ in a similar manner to that used in the paragraph after (2.17), such that for all t less than t_o in absolute value $|ch(tA\Sigma)| < 1$; using Lemmas 2.5 and 2.6, we obtain

$$\phi(t) = t\underset{\sim}{\mu}' \sum_{j=0}^{\infty} t^j 2^j (A\Sigma)^j A\underset{\sim}{\mu} + \frac{1}{2} \sum_{j=1}^{\infty} t^j 2^j \frac{\mathrm{tr}(A\Sigma)^j}{j}$$

$$= \sum_{j=1}^{\infty} t^j 2^{j-1} \underset{\sim}{\mu}'(A\Sigma)^{j-1}A\underset{\sim}{\mu} + \sum_{j=1}^{\infty} t^j 2^{j-1} \frac{\mathrm{tr}(A\Sigma)^j}{j} \quad ,$$

which upon simplification gives (2.26).

This result was obtained by Khatri (1963) for the singular case and again by Rohde, Urquhart, and Searle (1966). In his 1962 paper, Khatri obtained the joint cumulant generating function of two nonhomogeneous quadratic forms in nonsingular normal variates. Dieulefait (1951) and, apparently independently, Lancaster (1954) obtained the cumulant generating function for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, I)$.

COROLLARY 2.3: *The* j*th cumulant of the quadratic form* $\underset{\sim}{x}'A\underset{\sim}{x}$, *with* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, Σ *positive semidefinite, is*

$$K_j(\underset{\sim}{x}'A\underset{\sim}{x}) = j!2^{j-1}[\underset{\sim}{\mu}'(A\Sigma)^{j-1}A\underset{\sim}{\mu} + \mathrm{tr}\,\frac{(A\Sigma)^j}{j}\,] \,. \tag{2.27}$$

Proof: The right-hand side of (2.27) is the coefficient of $t^j/j!$ in (2.26).

COROLLARY 2.4: *The first four cumulants of* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *are*

$$K_1(\underset{\sim}{x}'A\underset{\sim}{x}) = E(\underset{\sim}{x}'A\underset{\sim}{x}) = \underset{\sim}{\mu}'A\underset{\sim}{\mu} + \mathrm{tr}A\Sigma \tag{2.28}$$

(valid also if $\underset{\sim}{x}$ *is not normal),*

$$K_2(\underset{\sim}{x}'A\underset{\sim}{x}) = V(\underset{\sim}{x}'A\underset{\sim}{x}) = 4\underset{\sim}{\mu}'A\Sigma A\underset{\sim}{\mu} + 2\mathrm{tr}(A\Sigma)^2 \,, \tag{2.29}$$

where $V(\cdot)$ *indicates variance,*

$$K_3(\underset{\sim}{x}'A\underset{\sim}{x}) = 24\underset{\sim}{\mu}'A\Sigma A\Sigma A\underset{\sim}{\mu} + 8\mathrm{tr}(A\Sigma)^3, \tag{2.30}$$

$$K_4(\underset{\sim}{x}'A\underset{\sim}{x}) = 192\underset{\sim}{\mu}'A\Sigma A\Sigma A\Sigma A\underset{\sim}{\mu} + 48\mathrm{tr}(A\Sigma)^4 \,. \tag{2.31}$$

The first four moments about the origin when $\underset{\sim}{\mu} = \underset{\sim}{0}$ *and* $\Sigma = I$ *are*

$$\mu_1(\underset{\sim}{x}'A\underset{\sim}{x}) = \mathrm{tr}A \tag{2.32}$$

$$\mu_2(\underset{\sim}{x}'A\underset{\sim}{x}) = 2\mathrm{tr}A^2 + (\mathrm{tr}A)^2 \tag{2.33}$$

$$\mu_3(\underset{\sim}{x}'A\underset{\sim}{x}) = 8\mathrm{tr}A^3 + 6\mathrm{tr}A^2 + (\mathrm{tr}A)^3 \tag{2.34}$$

$$\mu_4(\underset{\sim}{x}'A\underset{\sim}{x}) = 48\mathrm{tr}A^4 + 32\mathrm{tr}A^3\mathrm{tr}A + 12(\mathrm{tr}A^2)^2 + 12\mathrm{tr}A^2(\mathrm{tr}A)^2 + \mathrm{tr}A^4 . \tag{2.35}$$

Proof: (2.28)-(2.31) are obtained by direct substitution in (2.27). (2.32)-(2.35) are obtained by substituting the given conditions in the reversal formulas

$$\mu_1(\underset{\sim}{x}'A\underset{\sim}{x}) = K_1(\underset{\sim}{x}'A\underset{\sim}{x}) \tag{2.36}$$

$$\mu_2(\underset{\sim}{x}'A\underset{\sim}{x}) = K_2(\underset{\sim}{x}'A\underset{\sim}{x}) + K_1^2(\underset{\sim}{x}'A\underset{\sim}{x}) \tag{2.37}$$

$$\mu_3(\underset{\sim}{x}'A\underset{\sim}{x}) = K_3(\underset{\sim}{x}'A\underset{\sim}{x}) + 3K_2(\underset{\sim}{x}'A\underset{\sim}{x})K_1(\underset{\sim}{x}'A\underset{\sim}{x}) + K_1^3(\underset{\sim}{x}'A\underset{\sim}{x}) \tag{2.38}$$

$$\mu_4(\underset{\sim}{x}'A\underset{\sim}{x}) = K_4(\underset{\sim}{x}'A\underset{\sim}{x}) + 4K_3(\underset{\sim}{x}'A\underset{\sim}{x})K_1(\underset{\sim}{x}'A\underset{\sim}{x}) + 3K_2^2(\underset{\sim}{x}'A\underset{\sim}{x}) + 6K_2(\underset{\sim}{x}'A\underset{\sim}{x})K_1^2(\underset{\sim}{x}'A\underset{\sim}{x}) + K_1^4(\underset{\sim}{x}'A\underset{\sim}{x}) , \tag{2.39}$$

which are given, for example, in Kendall and Stuart (1969), Vol. I, pp. 68-71.

COROLLARY 2.5: *The cumulant generating function of the nonhomogeneous quadratic form* $Q = \underset{\sim}{x}'A\underset{\sim}{x} + 2\underset{\sim}{b}'\underset{\sim}{x} + c$, *where* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, Σ *positive semidefinite, is*

$$\phi(t) = t(\underset{\sim}{\mu}'A\underset{\sim}{\mu}+2\underset{\sim}{b}'\underset{\sim}{\mu}+c+\mathrm{tr}A\Sigma) + \sum_{j=2}^{\infty} \frac{t^j}{j!}[j!2^{j-1}((\underset{\sim}{\mu}'A+\underset{\sim}{b}')(\Sigma A)^{j-2}(\Sigma A\underset{\sim}{\mu}+\Sigma\underset{\sim}{b}) + \frac{\mathrm{tr}(A\Sigma)^j}{j})] \tag{2.40}$$

and the cumulants are:

$$K_1(Q) = \underset{\sim}{\mu}'A\underset{\sim}{\mu} + 2\underset{\sim}{b}'\underset{\sim}{\mu} + c + \mathrm{tr}A\Sigma \tag{2.41}$$

$$K_j(Q) = j!2^{j-1}[(\underset{\sim}{\mu}'A+\underset{\sim}{b}')(\Sigma A)^{j-2}(\Sigma A\underset{\sim}{\mu}+\Sigma\underset{\sim}{b}) + \frac{\mathrm{tr}(A\Sigma)^j}{j}],\ j > 1. \tag{2.42}$$

Proof: Using the same substitution in (2.27) as in Corollary 2.1 gives (2.40) after some simplification. The first cumulant is the coefficient of t in (2.40) which gives (2.41). The j*th* cumulant for $j > 1$ is the coefficient of $t^j/j!$ in (2.40) which gives (2.42).

The cumulant generating function for a bilinear form may be obtained in a similar manner to Corollary 2.2. However, the resultant expression is awkward.

CHAPTER III

Chi-Squaredness

Quadratic forms in or, equivalently, linear combinations of squares of normal variates occur throughout statistics. This is particularly so in the theory of regression and analysis of variance, as well as in time series analysis. The determination of the exact distribution of such quadratic forms is, therefore, of some importance. Although the cumulants and moments of all orders are available in closed form for any quadratic form in normal variables, the density or distribution function is available in closed form for surprisingly few. Amongst those for which the density function is known, the predominant ones are the central and noncentral chi-square and scalar multiples thereof.

In this chapter, necessary and sufficient conditions for a quadratic form in (possibly singular) normal variates to follow a central or noncentral chi-square distribution are established. Also, equivalent formulations of these conditions are discussed. As corollaries, simplified versions of these conditions are obtained for particular quadratic forms or covariance structures. Also, necessary and sufficient conditions for a nonhomogeneous quadratic form to follow a chi-square distribution are derived.

If $x_1,x_2,\ldots,x_r$ are independent and identically distributed $N_1(0,1)$, then $S = \sum_{i=1}^{r} x_i^2$ is said to follow a chi-square distribution with r degrees of freedom. This is denoted $S \sim \chi_r^2$. If the means of the x_i are not all zero, i.e., $E(x_i) = \mu_i$, $i=1,2,\ldots,r$, then S is said to follow a noncentral chi-square distribution with r degrees of freedom and noncentrality parameter $\delta^2 = \sum_{i=1}^{r} \mu_i^2$. This is denoted $S \sim \chi_r^2(\delta^2)$. The distribution is central when $\delta^2 = 0$ which occurs if and only if $E(x_i) = 0$ for all $i=1,2,\ldots,r$. It follows that

$$E[\chi_r^2(\delta^2)] = \delta^2 + r \; . \tag{3.1}$$

S may be written as $S = \underset{\sim}{x}'\underset{\sim}{x}$, where $\underset{\sim}{x} \sim N_r(\underset{\sim}{0},I)$ or $N_r(\underset{\sim}{\mu},I)$. From Theorem 2.1 it then follows that

$$E(e^{it\chi_r^2}) = (1-2it)^{-r/2} \tag{3.2}$$

and

$$E(e^{it\chi_r^2(\delta^2)}) = \frac{e^{it\delta^2/(1-2it)}}{(1-2it)^{r/2}} \; . \tag{3.3}$$

The fundamental result of this chapter is:

THEOREM 3.1: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, $\underset{\sim}{\mu}$ *not necessarily* $\underset{\sim}{0}$, *and* Σ *is positive semidefinite, then a set of necessary and sufficient conditions for* $\underset{\sim}{x}'A\underset{\sim}{x} \sim \chi_r^2(\delta^2)$ *is:*

$$\Sigma A \Sigma A \Sigma = \Sigma A \Sigma \tag{3.4}$$

$$\underset{\sim}{\mu}'(A\Sigma)^2 = \underset{\sim}{\mu}'A\Sigma \tag{3.5}$$

$$\underset{\sim}{\mu}'A\Sigma A\underset{\sim}{\mu} = \underset{\sim}{\mu}'A\underset{\sim}{\mu} \; , \tag{3.6}$$

and then

$$\mathrm{tr}(A\Sigma) = r \; , \quad \underset{\sim}{\mu}'A\underset{\sim}{\mu} = \delta^2 \; . \tag{3.7}$$

We note that (3.4), (3.5) and (3.6) may be compressed into

$$\begin{pmatrix} \Sigma \\ \mu' \end{pmatrix} A\Sigma A(\Sigma,\mu) = \begin{pmatrix} \Sigma \\ \mu' \end{pmatrix} A(\Sigma,\mu) .$$

<u>Proof:</u> A necessary condition for $x'Ax$ to follow a $\chi_r^2(\delta^2)$ distribution is that the moment generating function of $x'Ax$,

$$\phi_1(t) = \frac{e^{\mu'(I-2tA\Sigma)^{-1}A\mu}}{|I-2tA\Sigma|^{1/2}} \tag{3.8}$$

equals the moment generating function of $\chi_r^2(\delta^2)$,

$$\phi_2(t;r,\delta^2) = \frac{e^{t\delta^2/(1-2t)}}{(1-2t)^{r/2}} , \tag{3.9}$$

for some positive integer r and a real constant $\delta^2 \geq 0$. The function $\phi_1(t)$ is continuous for all real t except for a finite number of values. These values are at the points where $I-2tA\Sigma$ is singular and hence would have a determinant of zero. They occur when the characteristic roots of $I-2tA\Sigma$ (which may be denoted by $1-2t\lambda_j$, where $\lambda_1,\lambda_2,\ldots,\lambda_p$ are the characteristic roots of $A\Sigma$) vanish; i.e., when $1-2t\lambda_j = 0$ (or $t = (2\lambda_j)^{-1}$ for $\lambda_j \neq 0$). On the other hand, $\phi_2(t;r,\delta^2)$ is continuous with a single discontinuity at $t = 1/2$. As both functions have the same points of discontinuity, $(2\lambda_j)^{-1} = 1/2$ or $\lambda_j = 1$ for all nonzero λ_j. Thus r must equal the number of these nonzero roots and $r = \sum_{j=1}^{p} \lambda_j = tr(A\Sigma)$, which is the first part of (3.7). Let m be the rank of Σ and let $\Sigma = TT'$ where T is a real $p\times m$ matrix. Let $ch(A\Sigma)$ denote the characteristic roots of $A\Sigma$. Then the nonzero $ch(A\Sigma) = ch(ATT') = ch(T'AT)$. Thus, $T'AT$ must also have r roots equal to 1 and $m-r$ roots equal to zero. As $T'AT$ is symmetric, it must be idempotent of rank r. Therefore,

$$T'ATT'AT = T'AT \ . \tag{3.10}$$

Premultiplying (3.10) by T and postmultiplying by T' gives (3.4). Now, in order to use Lemma 2.5 to expand $(I-2tA\Sigma)^{-1}$, the characteristic roots of $2tA\Sigma$ must be less than 1 in absolute value. By selecting $t_o = \min(|(2\lambda_j)^{-1}|)$ for $j=1,2,\ldots,p$ and $\lambda_j \neq 0$, it follows that $|ch(2tA\Sigma)| < 1$ for all t such that $-t_o < t < t_o$. Using Lemma 2.5 gives

$$\mu'(I-2tA\Sigma)^{-1}A\mu = \mu' \sum_{k=0}^{\infty} (2t)^k (A\Sigma)^k A\mu \ . \tag{3.11}$$

As the exponent of (3.8) must equal the exponent of (3.9)

$$\mu' \sum_{k=0}^{\infty} (2t)^k (A\Sigma)^k A\mu = \delta^2/(1-2t) \ . \tag{3.12}$$

Simplifying gives

$$\mu'A\mu + \sum_{k=1}^{\infty} \mu'(2t)^k [(A\Sigma)^k - (A\Sigma)^{k-1}]A\mu = \delta^2 \ . \tag{3.13}$$

By (3.4), which we have already shown to be a necessary condition,

$$(A\Sigma)^3 = (A\Sigma)^2 \tag{3.14}$$

so that (3.13) reduces to

$$\mu'A\mu + 2t[\mu'A\Sigma A\mu - \mu'A\mu] + (2t)^2[\mu'A\Sigma A\Sigma A\mu - \mu'A\Sigma A\mu] = \delta^2 \tag{3.15}$$

for all t such that $-t_o < t < t_o$. As (3.15) holds for infinitely many values of t the coefficients of t^j are equal for $j=0,1,\ldots$, and we get $\mu'A\Sigma A\mu = \mu'A\mu$ or (3.6). Also

$$\mu' A\Sigma A\Sigma A\mu = \mu' A\Sigma A\mu \ . \tag{3.16}$$

Combining (3.6) and (3.16) gives

$$(\mu' A - \mu' A\Sigma A)\Sigma(A\Sigma A\mu - A\mu) = 0 \ . \tag{3.17}$$

Since Σ is positive semidefinite, (3.17) implies $(\mu' A - \mu' A\Sigma A)\Sigma = 0'$ and (3.5) is established. Using (2.28) and (3.1), we note that

$$E(x' Ax) = \mu' A\mu + tr(A\Sigma) = \delta^2 + r = E[\chi_r^2(\delta^2)] \ . \tag{3.18}$$

From (3.15), it is clear that $\mu' A\mu = \delta^2$; therefore from (3.18) we see that $r = tr(A\Sigma)$. Thus, necessity has been proven. To prove sufficiency, (3.4) - (3.6) are assumed and (3.8) must be shown to equal (3.9); i.e., the moment generating functions must be equal.

From (3.14) all the nonzero characteristic roots of $A\Sigma$ are equal to 1. If there are r of these then $r = tr(A\Sigma)$ and

$$|I - 2tA\Sigma|^{1/2} = \prod_{j=1}^{p} (1 - 2t\lambda_j)^{1/2} = (1-2t)^{r/2} \ . \tag{3.19}$$

From (3.14), (3.5) and (3.6), $\mu'(A\Sigma)^k A\mu = \mu' A\mu$ for $k=0,1,\ldots$. Thus,

$$\begin{aligned} \mu'(I - 2tA\Sigma)^{-1}A\mu &= \sum_{k=0}^{\infty} (2t)^k \mu'(A\Sigma)^k A\mu = \delta^2 \sum_{k=0}^{\infty} (2t)^k \\ &= \delta^2/(1-2t) \ , \end{aligned} \tag{3.20}$$

with $\delta^2 = \mu' A\mu$ and $-\frac{1}{2} < t < \frac{1}{2}$. (3.19) and (3.20) show that the moment generating functions are equal and so the theorem is proven.

COROLLARY 3.1: *If* $x \sim N_p(\mu, \Sigma)$ *and* $x' Ax$ *follow a chi-square dis-*

tribution, then the distribution is central if and only if

$$\mu' A \mu = 0 \quad \textit{and} \quad \Sigma A \mu = 0 \ . \tag{3.21}$$

<u>Proof</u>: If the chi-square distribution is central $\delta^2 = 0$ and $\mu' A \mu = 0$. Also $\mu' A \mu = \mu' A \Sigma A \mu = (T' A \mu)' T' A \mu = 0$ which implies $T' A \mu = 0$ or $\Sigma A \mu = 0$. Conversely, if $\mu' A \mu = 0$ and $\Sigma A \mu = 0$ then $\delta^2 = 0$ and the distribution is central.

Ogasawara and Takahashi (1951), Rao (1962), Khatri (1963), Rayner and Livingstone (1965), and Mäkeläinen (1966) all, apparently independently, obtained the above results.

The following theorem states an equivalent set of necessary and sufficient conditions for the quadratic form $x' A x$ to follow a $\chi_r^2(\delta^2)$ distribution.

THEOREM 3.2: *If* $x \sim N_p(\mu, \Sigma)$, *with* Σ *positive semidefinite, then a set of necessary and sufficient conditions for* $x' A x$ *to be distributed as* $\chi_r^2(\delta^2)$ *is*

$$\mu' (A\Sigma)^{j-1} A \mu = \delta^2 \ ; \quad j = 1,2,3,4 \tag{3.22}$$

and

$$\mathrm{tr}(A\Sigma)^4 = \mathrm{tr}(A\Sigma)^3 = \mathrm{tr}(A\Sigma)^2 = \mathrm{tr}(A\Sigma) = r \ . \tag{3.23a}$$

If A *is positive semidefinite then* (3.23a) *may be replaced by*

$$\mathrm{tr}(A\Sigma)^3 = \mathrm{tr}(A\Sigma)^2 = \mathrm{tr}(A\Sigma) = r \ . \tag{3.23b}$$

<u>Proof</u>: We must show that (3.22) and (3.23) are equivalent to (3.4), (3.5) and (3.6). Clearly (3.4), (3.5) and (3.6) imply (3.22) and (3.23).

Now from (3.22) we obtain

$$\underset{\sim}{\mu}'A\Sigma A\underset{\sim}{\mu} = \underset{\sim}{\mu}'A\Sigma A\Sigma A\underset{\sim}{\mu} = \underset{\sim}{\mu}'A\Sigma A\Sigma A\Sigma A\underset{\sim}{\mu} , \tag{3.24}$$

or, equivalently, by writing $\Sigma = TT'$

$$\underset{\sim}{\mu}'(A\Sigma AT-AT)(A\Sigma AT-AT)'\underset{\sim}{\mu} = 0 . \tag{3.25}$$

This implies

$$\underset{\sim}{\mu}'A\Sigma AT = \underset{\sim}{\mu}'AT , \tag{3.26}$$

which by postmultiplying by T' gives (3.5). Note that (3.22) also implies (3.6). Let $\lambda_1,\ldots,\lambda_p$ be the characteristic roots of $A\Sigma$ and hence of $T'AT$. Since $T'AT$ is symmetric, all the λ_i are real and (3.23a) becomes

$$\sum_{i=1}^{p} \lambda_i^4 = \sum_{i=1}^{p} \lambda_i^3 = \sum_{i=1}^{p} \lambda_i^2 = \sum_{i=1}^{p} \lambda_i , \tag{3.27}$$

which implies

$$\sum_{i=1}^{p} (\lambda_i^4-2\lambda_i^3+\lambda_i^2) = \sum_{i=1}^{p} \lambda_i^2(\lambda_i-1)^2 = 0 . \tag{3.28}$$

But as each $\lambda_i^2(\lambda_i-1)^2 \geq 0$, the characteristic roots λ_i must either be equal to 0 or 1 and as $r = tr(A\Sigma)$ there must be r characteristic roots equal to 1. Thus $T'AT$ is idempotent and

$$T'ATT'AT = T'AT , \tag{3.29}$$

which is equivalent to (3.5). When A is positive semidefinite, we know that the λ_i must all be nonnegative since $T'AT$ is then also positive semidefinite. (3.23b) implies

$$\sum_{i=1}^{p} \lambda_i^3 = \sum_{i=1}^{p} \lambda_i^2 = \sum_{i=1}^{p} \lambda_i \tag{3.30}$$

or

$$\sum_{i=1}^{p} (\lambda_i^3 - 2\lambda_i^2 + \lambda_i) = \sum_{i=1}^{p} \lambda_i(\lambda_i - 1)^2 = 0 \ . \tag{3.31}$$

But each $\lambda_i(\lambda_i-1)^2 \geq 0$, therefore each λ_i must equal either 0 or 1. Again, as $r = \mathrm{tr}(A\Sigma)$, r of the λ_i equal 1 and the remainder 0. This implies that $T'AT$ is idempotent and we obtain (3.29) which is equivalent to (3.5). Thus Theorem 3.2 is proven.

The latter part of the above proof is due to Shanbhag (1970) extending a previous result of Good (1969). The following theorem is given as an exercise by Searle (1971).

THEOREM 3.2a: *If* $x \sim N_p(\mu, \Sigma)$, *with* Σ *positive semidefinite, then a necessary and sufficient condition for* $x'Ax$ *to be distributed as* $\chi_r^2(\delta^2)$ *is*

$$\mathrm{tr}(A\Sigma)^j + j\mu'(A\Sigma)^{j-1}A\mu = r + j\delta^2 \ ; \ \ j = 1,2,\ldots , \tag{3.32}$$

where $\delta^2 = \mu'A\mu$.

<u>Proof</u>: The cumulant generating function of $\chi_r^2(\delta^2)$, obtained from the logarithm of its moment generating function in (3.9) is

$$\phi(t) = t\delta^2/(1-2t) - r[\log(1-2t)]/2 \ . \tag{3.33}$$

Expanding the right-hand side as a power series in t, for $|t| < 1/2$, gives the jth cumulant as the coefficient of $t^j/j!$; that is,

$$K_j[\chi_r^2(\delta^2)] = 2^{j-1}(j-1)!(r+j\delta^2) \; ; \quad j = 1,2,\ldots \; . \tag{3.34}$$

The j*th* cumulant of $\underset{\sim}{x}'A\underset{\sim}{x}$, as given by (2.27), is

$$K_j[\underset{\sim}{x}'A\underset{\sim}{x}] = 2^{j-1}j![\underset{\sim}{\mu}'(A\Sigma)^{j-1}A\underset{\sim}{\mu}+\mathrm{tr}(A\Sigma)^j/j] \; ; \quad j = 1,2,\ldots \; . \tag{3.35}$$

Since the moments of the chi-square distribution determine that distribution uniquely (cf. Anderson, 1958, p. 172 and Rao, 1965, p. 86), it follows that a necessary and sufficient condition for $\underset{\sim}{x}'A\underset{\sim}{x}$ to be distributed $\chi_r^2(\delta^2)$ is that the right-hand side of (3.34) equals the right-hand side of (3.35) for $j=1,2,\ldots$. After simplification, this gives (3.32) as a necessary and sufficient condition for $\underset{\sim}{x}'A\underset{\sim}{x}$ to have a $\chi_r^2(\delta^2)$ distribution.

Theorems 3.2 and 3.2a show that (3.32) is equivalent to (3.22) and (3.23a), a result we have not been able to prove directly.

Another set of necessary and sufficient conditions for $\underset{\sim}{x}'A\underset{\sim}{x}$ to be distributed $\chi_r^2(\delta^2)$ is given by Rao and Mitra (1971). The set of conditions is (i) $\Sigma A\Sigma A\Sigma = \Sigma A\Sigma$, (ii) $\underset{\sim}{\mu}'A\Sigma A\underset{\sim}{\mu} = \underset{\sim}{\mu}'A\underset{\sim}{\mu}$ and (iii) $\Sigma A\underset{\sim}{\mu}$ belongs to the column space of $\Sigma A\Sigma$. The proof shows the equivalence of (iii) and $\Sigma A\Sigma A\underset{\sim}{\mu} = \Sigma A\underset{\sim}{\mu}$. Rohde, Urquhart and Searle (1966) knew of this result from personal communications with Rao and used it to prove the following. Let $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, with $\Sigma = TT'$, where T is a real $p\times m$ matrix and $m = \mathrm{rk}(\Sigma)$ and let $\underset{\sim}{y} \sim N_m(\underset{\sim}{0},I)$. Then a necessary and sufficient condition for $\underset{\sim}{x}'A\underset{\sim}{x}$ to have a $\chi_r^2(\delta^2)$ distribution is that there exists a real $m\times 1$ vector $\underset{\sim}{c}$ for which

$$\underset{\sim}{x}'A\underset{\sim}{x} = (\underset{\sim}{y}+\underset{\sim}{c})'T'AT(\underset{\sim}{y}+\underset{\sim}{c}) \; , \tag{3.36}$$

with T'AT idempotent.

We now present corollaries to Theorem 3.1 in which additional conditions imposed upon the covariance matrix and/or the matrix of the quadratic form simplify the necessary and sufficient conditions for chi-squaredness.

COROLLARY 3.2: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *and if* Σ *is positive definite, then* $\underset{\sim}{x}'A\underset{\sim}{x}$ *follows a* $\chi_r^2(\delta^2)$ *distribution if and only if*

$$A\Sigma A = A \ , \qquad (3.37)$$

and then $r = rk(A)$ *and* $\delta^2 = \underset{\sim}{\mu}'A\underset{\sim}{\mu}$. *The distribution is central if and only if* $A\underset{\sim}{\mu} = \underset{\sim}{0}$.

Proof: When Σ is positive definite, (3.37) readily implies (3.4), (3.5) and (3.6) and thus that $\underset{\sim}{x}'A\underset{\sim}{x}$ is distributed $\chi_r^2(\delta^2)$ with $r = tr(A\Sigma) = rk(A\Sigma) = rk(A)$ and $\delta^2 = \underset{\sim}{\mu}'A\underset{\sim}{\mu}$. When $\underset{\sim}{x}'A\underset{\sim}{x}$ is distributed as $\chi_r^2(\delta^2)$ then (3.4), (3.5) and (3.6) hold. If Σ is positive definite, (3.20) follows. Also, whereas $\Sigma A\underset{\sim}{\mu} = \underset{\sim}{0}$ and $\underset{\sim}{\mu}'A\underset{\sim}{\mu} = 0$ are necessary and sufficient for central chi-square, the positive definiteness of Σ reduce these conditions to $A\underset{\sim}{\mu} = \underset{\sim}{0}$.

Corollary 3.2 was given by Carpenter (1950), Graybill and Marsaglia (1957) and Rao (1965); when $\underset{\sim}{\mu} = \underset{\sim}{0}$, Ogawa (1946b) indicated that it was proven by Sakamoto in 1943. (Cf. Sakamoto, 1944b.)

We note that (3.37) implies that Σ is a generalized inverse of A (cf. Rao and Mitra, 1971); we write $\Sigma = g_1(A)$. We also see from (3.37) that A is positive semidefinite.

COROLLARY 3.3: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ *positive semidefinite and* $rk(\Sigma) = r \leq p$, *then a set of necessary and sufficient conditions for* $\underset{\sim}{x}'A\underset{\sim}{x}$ *to be distributed* $\chi_r^2(\delta^2)$ *is that*

$$A = g_1(\Sigma) \quad or \quad \Sigma A \Sigma = \Sigma \tag{3.38}$$

and

$$\underset{\sim}{\mu}'A\Sigma A\underset{\sim}{\mu} = \underset{\sim}{\mu}'A\underset{\sim}{\mu} , \tag{3.39}$$

where $\delta^2 = \underset{\sim}{\mu}'A\underset{\sim}{\mu}$. *The distribution is central if and only if* $\underset{\sim}{\mu}'A\underset{\sim}{\mu} = 0$ *or, equivalently,* $\Sigma A\underset{\sim}{\mu} = \underset{\sim}{0}$.

We note that this result differs from Theorem 3.1, in that here we specify that r, the degrees of freedom, must equal $rk(\Sigma)$.

<u>Proof</u>: Let $\Sigma = TT'$ where T is a $p \times r$ matrix. If $\underset{\sim}{x}'A\underset{\sim}{x}$ is distributed $\chi_r^2(\delta^2)$ then (3.4) implies that $T'AT$ is idempotent. As $r = tr(A\Sigma) = tr(T'AT) = rk(T'AT)$, $T'AT$ is nonsingular. Since the only nonsingular idempotent matrix is the identity, $T'AT = I$, and so $TT'ATT' = TT'$ or (3.38). Also, (3.6) is identical to (3.39). On the other hand, when (3.38) and (3.39) hold, (3.4), (3.5) and (3.6) follow and thus $\underset{\sim}{x}'A\underset{\sim}{x} \sim \chi_r^2(\delta^2)$, with $\delta^2 = \underset{\sim}{\mu}'A\underset{\sim}{\mu}$. The conditions for central chi-square remain unchanged.

This corollary was proven for $\underset{\sim}{\mu} = \underset{\sim}{0}$ by Khatri (1968) and by Zelen and Federer (1965). It is interesting to note that if $rk(\Sigma) = p$, then $\Sigma A\Sigma = \Sigma$ reduces to $A = \Sigma^{-1}$, a result proved differently by Bhat (1962).

COROLLARY 3.4: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\Sigma)$, Σ *positive semidefinite, then* $\underset{\sim}{x}'A\underset{\sim}{x}$

follows a χ_r^2 distribution *if and only if*

$$\Sigma A \Sigma A \Sigma = \Sigma A \Sigma \ , \tag{3.40}$$

and then $r = tr(A\Sigma)$. *If* Σ *is positive definite, then* (3.40) *reduces to* (3.37).

The proof follows directly from Theorem 3.1. Rayner and Livingstone (1965) claim to have had this result in 1955 and draw attention to an incorrect version in Rao (1962).

The following theorems and corollaries will try to relax condition (3.40) given certain additional conditions on the trace or rank of combinations of A and Σ. The motivation for this is to reduce the necessary and sufficient condition (3.40) for chi-squaredness. Indeed, with certain additional information, we do have simpler necessary conditions.

LEMMA 3.1: *If* Σ *is a* $p\times p$ *positive semidefinite matrix and* $A = A'$, *then*

(1) $$\Sigma A \Sigma A \Sigma = \Sigma A \Sigma \quad \Leftrightarrow \quad (\Sigma A)^3 = (\Sigma A)^2 \ , \tag{3.41}$$

(2) $$rk(A\Sigma) = rk(\Sigma A) = rk(A\Sigma A) \ , \tag{3.42}$$

(3) $$rk[(A\Sigma)^2] = rk(\Sigma A \Sigma) = rk[(\Sigma A)^2] \ . \tag{3.43}$$

<u>Proof</u>: As Σ is positive semidefinite, we may write $\Sigma = TT'$ where T is a $p\times s$ matrix and $s = rk(\Sigma)$.

(1) Clearly $\Sigma A \Sigma A \Sigma = \Sigma A \Sigma$ implies $(\Sigma A)^3 = (\Sigma A)^2$. Now $(\Sigma A)^3 = (\Sigma A)^2$ implies

$$\not{T}A\not{T}A\not{T}A\not{T}A\not{T} = \not{T}A\not{T}A\not{T}A\not{T} = \not{T}A\not{T}A\not{T} . \qquad (3.44)$$

This yields

$$(\not{T}A\not{T}AT-\not{T}AT)(\not{T}A\not{T}AT-\not{T}AT)' = 0 . \qquad (3.45)$$

which gives $\not{T}A\not{T}AT = \not{T}AT$ and so $\not{T}A\not{T}A\not{T} = \not{T}A\not{T}$.

(2) $rk(A\not{T}) = rk[(A\not{T})'] = rk(\not{T}A)$. In addition $rk(A\not{T}) = rk(ATT') = rk(AT) = rk[(AT)(AT)'] = rk(A\not{T}A)$.

(3) $rk(\not{T}A\not{T}) = rk(TT'ATT') = rk(T'AT) = rk[(T'AT)(T'AT)']$ $= rk(T'ATT'AT) = rk(TT'ATT'ATT') = rk(\not{T}A\not{T}A\not{T})$. But as $rk(\not{T}A\not{T})$ $\geq rk(\not{T}A\not{T}A) = rk(A\not{T}A\not{T}) \geq rk(\not{T}A\not{T}A\not{T})$ and $rk(\not{T}A\not{T}) = rk(\not{T}A\not{T}A\not{T})$, the inequality string collapses and (3.43) follows.

COROLLARY 3.5: *Let* $r_1 = rk(A\not{T})$ *and* $r_2 = rk[(A\not{T})^2]$. *If* $s = rk(\not{T})$ *and* $r = rk(A)$, *then*

$$s \geq r_1 \geq r_2 \quad \textit{and} \quad r \geq r_1 \geq r_2 . \qquad (3.46)$$

$$\textit{If} \quad s=r_2 \quad \textit{then} \quad s=r_1=r_2 \quad \textit{and if} \quad r=r_2, \quad \textit{then} \quad r=r_1=r_2 . \qquad (3.47)$$

<u>Proof:</u> $s = rk(\not{T}) \geq rk(A\not{T}) \geq rk[(A\not{T})^2]$ and $r = rk(A) \geq rk(A\not{T}) \geq rk[(A\not{T})^2]$. When the extremes are equal, equality holds throughout, i.e., (3.47).

THEOREM 3.3: (i) *If* A *is positive semidefinite then* $r_1 = r_2$ *where* $r_1 = rk(A\not{T})$ *and* $r_2 = rk[(A\not{T})^2]$. (ii) *If* $A\not{T}$ *is symmetric, then* $r_1 = r_2$.

<u>Proof:</u> (i) From Lemma 3.1 $r_2 = rk(\not{T}A\not{T})$. Since A is positive semi-

definite, we may write $A = FF'$, where F is a $p \times r$ matrix. Then $rk(ZAZ) = rk(ZFF'Z) = rk[(ZF)(ZF)'] = rk(ZF) = rk(ZFF') = rk(ZA) = r_1$.

(ii) $r_2 = rk[(AZ)^2] = rk[(AZ)(AZ)] = rk[(AZ)(AZ)'] = rk(AZ) = r_1$.

THEOREM 3.4: (i) $ZAZAZ = ZAZ \Rightarrow rk(ZAZ) = tr(AZ)$

(ii) $ZAZAZ = ZAZ$ *and* $rk(AZ) = rk[(AZ)^2] \Rightarrow AZAZ = AZ$

(iii) $ZAZAZ = ZAZ$ *and* $rk(A) = rk(AZ) \Rightarrow AZA = A$

(iv) $AZAZ = AZ$ *and* $rk(Z) = rk(AZ) \Rightarrow ZAZ = Z$.

We prove this theorem using the following lemma due to Styan (1971).

LEMMA 3.2: *If* $GHB = GHC$ *and* $rk(GH) = rk(H)$ *then* $HB = HC$.

Proof: We may write $H = KL'$ where K and L have full column rank. Then $rk(GH) = rk(H)$ implies $rk(GK) = r(K)$ and thus GK has full column rank. Hence $GHB = GHC$, or $GKL'B = GKL'C$, implies $L'B = L'C$ or $HB = HC$.

Proof of Theorem 3.4: (i) $ZAZAZ = ZAZ$ implies $T'ATT'AT = T'AT$, or $T'AT$ idempotent. Thus $rk(T'AT) = tr(T'AT)$. Since $rk(T'AT) = rk(ZAZ)$ and $tr(T'AT) = tr(AZ)$ we have $rk(ZAZ) = tr(AZ)$.

(ii) Since $rk(AZ) = rk[(AZ)^2]$, we have $rk(ZAZ) = rk(AZ)$. Applying Lemma 3.2 to $ZAZAZ = ZAZ$ with $G = Z$, $H = AZ$, $B = AZ$, and $C = I$ we get $AZAZ = AZ$.

(iii) Since $rk(A) = rk(AZ) = rk(ZA)$, using Lemma 3.2 on $ZAZAZ$. with $G = Z$, $H = A$, $B = Z$, and $C = AZ$, gives $AZAZ = AZ$. Again by transposing and reapplying the lemma, we obtain $A\Sigma A = A$.

(iv) Since $\mathrm{rk}(\Sigma) = \mathrm{rk}(A\Sigma)$, applying the above lemma to $A\Sigma A\Sigma = A\Sigma$ gives $\Sigma A\Sigma = \Sigma$, as required.

COROLLARY 3.6: (i) *If* $\Sigma A\Sigma A\Sigma = \Sigma A\Sigma$ *and* $\mathrm{rk}(A) = \mathrm{rk}[(A\Sigma)^2]$ *then* $A\Sigma A = A$. (ii) *If* $A\Sigma A\Sigma = A\Sigma$ *and* $\mathrm{rk}(\Sigma) = \mathrm{rk}[(A\Sigma)^2]$ *then* $\Sigma A\Sigma = \Sigma$.

<u>Proof</u>: Using Corollary 3.5, $\mathrm{rk}(A) = \mathrm{rk}[(A\Sigma)^2]$ implies $\mathrm{rk}(A) = \mathrm{rk}(A\Sigma)$. Also, $\mathrm{rk}(\Sigma) = \mathrm{rk}[(A\Sigma)^2]$ implies $\mathrm{rk}(\Sigma) = \mathrm{rk}(A\Sigma)$. Using Lemma 3.2 establishes (i) and (ii).

Parts of the above theorems and corollaries from Theorem 3.3 onwards occur throughout the literature. See, for example, Rayner and Livingstone (1965), Shanbhag (1968), Good (1969), Styan (1970), Rayner and Nevin (1970), and Rao and Mitra (1971).

When $\mathbf{x} \sim N_p(\mathbf{0},\Sigma)$, with Σ positive semidefinite, the necessary and sufficient conditions for $\mathbf{x}'A\mathbf{x}$ to be distributed χ^2_r reduce from (3.4), (3.5) and (3.6) of Theorem 3.1 to

$$\Sigma A\Sigma A\Sigma = \Sigma A\Sigma \,, \tag{3.48}$$

and then $r = \mathrm{tr}(A\Sigma)$ as in Corollary 3.4. Shanbhag (1970) obtained an equivalent formulation of the above results as follows:

COROLLARY 3.7: *Let* $\mathbf{x} \sim N_p(\mathbf{0},\Sigma)$, *with* Σ *positive semidefinite. Then* $\mathbf{x}'A\mathbf{x}$ *follows a* χ^2_r *distribution if and only if*

$$\mathrm{tr}(A\Sigma)^4 = \mathrm{tr}(A\Sigma)^3 = \mathrm{tr}(A\Sigma)^2 = \mathrm{tr}(A\Sigma) = r \,. \tag{3.49}$$

If A *is positive semidefinite then* $\mathbf{x}'A\mathbf{x} \sim \chi^2_r$ *if and only if*

$$\mathrm{tr}(A\Sigma)^3 = \mathrm{tr}(A\Sigma)^2 = \mathrm{tr}(A\Sigma) = r \ . \tag{3.50}$$

The proof is immediate from Theorem 3.2.

The next two corollaries, whose proofs are also immediate (and therefore omitted), are included to give historical perspective.

COROLLARY 3.8a: [Cochran (1934), Craig (1943)] *If* $x \sim N_p(0,I)$, *then* $x'Ax$ *follows a chi-square distribution if and only if all the nonzero characteristic roots of* A *are* 1 *or, equivalently,* $A^2 = A$.

COROLLARY 3.8b: [Carpenter (1950)] *If* $x \sim N_p(\mu,\sigma^2 I)$, *then* $x'Ax$ *follows a chi-square distribution* $\chi_r^2(\delta^2)$ *if and only if* $\sigma^2A^2 = A$, *and then* $r = \mathrm{rk}(A)$ *and* $\delta^2 = \mu'A\mu$.

The necessary and sufficient conditions for a nonhomogeneous quadratic form in normal variables to be distributed $\chi_r^2(\delta^2)$ are now obtained using the same technique as in Corollary 2.1.

COROLLARY 3.9: *If* $x \sim N_p(\mu,\Sigma)$, Σ *positive semidefinite, then* $x'Ax + 2b'x + c$ *is distributed* $\chi_r^2(\delta^2)$ *if and only if*

$$\Sigma A\Sigma A\Sigma = \Sigma A\Sigma \tag{3.51}$$

$$(A\mu+b)'\Sigma A\Sigma = (A\mu+b)'\Sigma \tag{3.52}$$

$$(A\mu+b)'\Sigma(A\mu+b) = \mu'A\mu + 2b'\mu + c \ , \tag{3.53}$$

and then $r = \mathrm{tr}(A\Sigma)$ *and* $\delta^2 = \mu'A\mu + 2b'\mu + c$. *The distribution is central if and only if*

$$(A\mu+b)'\Sigma = 0' \ . \tag{3.54}$$

<u>Proof</u>: *If* $\underset{\sim}{x}_o = \begin{pmatrix} \underset{\sim}{x} \\ 1 \end{pmatrix}$, then $\underset{\sim}{x}_o \sim N_{p+1}(\underset{\sim}{\mu}_o, \Sigma_o)$, where $\underset{\sim}{\mu}_o = \begin{pmatrix} \underset{\sim}{\mu} \\ 1 \end{pmatrix}$ and $\Sigma_o = \begin{pmatrix} \Sigma & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{pmatrix}$. If $A_o = \begin{pmatrix} A & \underset{\sim}{b} \\ \underset{\sim}{b}' & c \end{pmatrix}$, then $\underset{\sim}{x}_o' A_o \underset{\sim}{x}_o = \underset{\sim}{x}' A \underset{\sim}{x} + 2\underset{\sim}{b}'\underset{\sim}{x} + c$. Moreover, $\underset{\sim}{x}_o' A_o \underset{\sim}{x}_o$ follows a $\chi_r^2(\delta^2)$ distribution if and only if (by Theorem 3.1)

$$\Sigma_o A_o \Sigma_o A_o \Sigma_o = \Sigma_o A_o \Sigma_o \tag{3.55}$$

$$\underset{\sim}{\mu}_o' A_o \Sigma_o = \underset{\sim}{\mu}_o' (A_o \Sigma_o)^2 \tag{3.56}$$

$$\underset{\sim}{\mu}_o' A_o \underset{\sim}{\mu}_o = \underset{\sim}{\mu}_o' A_o \Sigma_o A_o \underset{\sim}{\mu}_o , \tag{3.57}$$

and then $r = tr(A_o \Sigma_o)$ and $\delta^2 = \underset{\sim}{\mu}_o' A_o \underset{\sim}{\mu}_o$. Substituting for Σ_o, A_o, and $\underset{\sim}{\mu}_o$ in (3.55) gives (3.51), in (3.56) gives (3.52) and in (3.57) gives (3.53). The distribution is central if and only if $(A\underset{\sim}{\mu}+\underset{\sim}{b})'\Sigma(A\underset{\sim}{\mu}+\underset{\sim}{b}) = 0$. This holds if and only if $(A\underset{\sim}{\mu}+\underset{\sim}{b})'T = \underset{\sim}{0}'$ or, equivalently, $(A\underset{\sim}{\mu}+\underset{\sim}{b})'\Sigma = \underset{\sim}{0}'$.

This result was obtained by Khatri (1963). Ogasawara and Takahashi (1951) found the necessary and sufficient conditions for a nonhomogeneous quadratic form to follow a central chi-square distribution. Khatri (1962) considered $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, I)$.

We now examine the necessary and sufficient conditions for a bilinear form to be distributed as $\chi_r^2(\delta^2)$.

THEOREM 3.5: *Let* $\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \end{pmatrix} \sim N_{p+q}(\underset{\sim}{\mu}, \Sigma)$, *where* $\underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$ *and* $\underset{\sim}{x}_1 \sim N_p(\underset{\sim}{\mu}_1, \Sigma_{11})$ *and* $\underset{\sim}{x}_2 \sim N_q(\underset{\sim}{\mu}_2, \Sigma_{22})$. *Then* $\underset{\sim}{x}_1' A_o \underset{\sim}{x}_2$, *where* A_o

is any p×q *matrix, never follows a* $\chi_r^2(\delta^2)$ *distribution.*

<u>Proof</u>: Let $A = \frac{1}{2}\begin{pmatrix} 0 & A_o \\ A_o' & 0 \end{pmatrix}$. Then $\underset{\sim}{x}'A\underset{\sim}{x} = \underset{\sim}{x}_1'A_o\underset{\sim}{x}_2$. Were $\underset{\sim}{x}_1'A_o\underset{\sim}{x}_2$ or equivalently $\underset{\sim}{x}'A\underset{\sim}{x}$ distributed as $\chi_r^2(\delta^2)$, then

$$r = \mathrm{tr}(A\Sigma)$$

$$= \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 0 & A_o \\ A_o' & 0 \end{pmatrix}\begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}\right]$$

$$= \frac{1}{2}\,\mathrm{tr}\begin{pmatrix} 0 & A_o\Sigma_{22} \\ A_o'\Sigma_{11} & 0 \end{pmatrix} = 0 .$$

Hence, $\underset{\sim}{x}_1'A_o\underset{\sim}{x}_2$ cannot follow a chi-square distribution.

We now derive necessary and sufficient conditions for an arbitrary bilinear form to follow a $\chi_r^2(\delta^2)$ distribution.

THEOREM 3.6: *Let* $\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \end{pmatrix} \sim N_{p+q}(\underset{\sim}{\mu}, \Sigma)$, *where* $\underset{\sim}{x}_1$ *is* p×1, $\underset{\sim}{x}_2$ *is* q×1, $\underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, *with* Σ *positive semidefinite. Let* A_o *be a* p×q *matrix; then a set of necessary and sufficient conditions for* $\underset{\sim}{x}_1'A_o\underset{\sim}{x}_2$ *to be distributed* $\chi_r^2(\delta^2)$ *is* (3.4), (3.5) *and* (3.6) *with* $A = \frac{1}{2}\begin{pmatrix} 0 & A_o \\ A_o' & 0 \end{pmatrix}$ *and* $\underset{\sim}{\mu}$ *and* Σ *as above. Then* $r = \mathrm{tr}(A_o\Sigma_{21})$ *and* $\delta^2 = \underset{\sim}{\mu}_1'A_o\underset{\sim}{\mu}_2$.

<u>Proof</u>: Since $\underset{\sim}{x}'A\underset{\sim}{x} = \underset{\sim}{x}_1'A_o\underset{\sim}{x}_2$, conditions (3.4), (3.5) and (3.6), which are necessary and sufficient for $\underset{\sim}{x}'A\underset{\sim}{x}$ to be distributed $\chi_r^2(\delta^2)$, are

equally necessary and sufficient for $\underset{\sim}{x}_1' A_o \underset{\sim}{x}_2$ to have a $\chi_r^2(\delta^2)$ distribution. In addition,

$$r = \mathrm{tr}(A\Sigma)$$

$$= \mathrm{tr}\,\frac{1}{2}\begin{pmatrix} 0 & A_o \\ A_o' & 0 \end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

$$= \mathrm{tr}\,\frac{1}{2}\begin{pmatrix} A_o\Sigma_{21} & A_o\Sigma_{22} \\ A_o'\Sigma_{11} & A_o'\Sigma_{12} \end{pmatrix}$$

$$= \frac{1}{2}\,\mathrm{tr}(A_o\Sigma_{21}) + \frac{1}{2}\,\mathrm{tr}(A_o'\Sigma_{12}) = \mathrm{tr}(A_o\Sigma_{21}) \;,$$

and $\delta^2 = \underset{\sim}{\mu}' A \underset{\sim}{\mu} = \underset{\sim}{\mu}_1' A_o \underset{\sim}{\mu}_2$.

CHAPTER IV

Independence

In this chapter we present necessary and sufficient conditions for two quadratic forms in noncentral, possibly singular normal variables to be independently distributed. For corollaries, we obtain conditions for the independence of (i) two nonhomogeneous quadratic forms, (ii) two bilinear forms and (iii) a quadratic form and a set of linear forms. We also study the special cases of central variables, of a positive definite covariance matrix, and of positive semidefinite quadratic forms.

Independence of random variables is very useful to know. When two random variables are independent, their joint distribution function is the product of the two marginal distribution functions. Also, the ratios of certain random variables follow well-known distributions if the numerator and denominator random variables are independent. Examples of this are (i) the ratio of a standard normal variable to the square root of an independently distributed central chi-square variable divided by its degrees of freedom follows Student's t distribution, and (ii) the ratio of two independent chi-square variables, each divided by its degrees of freedom, follows the F distribution. These facts are extremely useful in the analysis of linear models; e.g., regression and the analysis of variance. It should be stressed, however, that the conditions for the independence of two quadratic forms do not in general depend on the individual distributions of the quadratic forms (e.g., whether or not they are chi-square).

THEOREM 4.1: *If* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ *positive semidefinite, and* A,B *are* p×p *symmetric matrices then a set of necessary and sufficient conditions for the quadratic forms* $\underset{\sim}{x}'A\underset{\sim}{x}$ *and* $\underset{\sim}{x}'B\underset{\sim}{x}$ *to be independent is*

$$\Sigma A \Sigma B \Sigma = 0 , \tag{4.1}$$

$$\Sigma B \Sigma A \underset{\sim}{\mu} = \underset{\sim}{0} , \tag{4.2}$$

$$\Sigma A \Sigma B \underset{\sim}{\mu} = \underset{\sim}{0} , \tag{4.3}$$

$$\underset{\sim}{\mu}' A \Sigma B \underset{\sim}{\mu} = 0 , \tag{4.4}$$

or

$$\begin{pmatrix} \Sigma \\ \underset{\sim}{\mu}' \end{pmatrix} A \Sigma B (\Sigma, \underset{\sim}{\mu}) = 0 . \tag{4.4a}$$

We will use the following lemma to help prove Theorem 4.1.

LEMMA 4.1: *If* A *and* B *are symmetric matrices then*

$$|I-sA|.|I-tB| = |I-sA-tB| \quad \textit{for all real} \quad s,t \tag{4.5}$$

if and only if

$$AB = 0 . \tag{4.6}$$

<u>Proof</u>: Since $|I-sA|.|I-tB| = |I-sA-tB+stAB|$ for all s and t, it is clear that (4.6) implies (4.5).

For all s,t sufficiently small such that

$$|ch(sA)| < 1 , \quad |ch(tB)| < 1 , \tag{4.7}$$

we may take logarithms of both sides of (4.5) and obtain

$$\log(|I-sA|) + \log(|I-tB|) = \log(|I-sA-tB|) . \tag{4.8}$$

We may apply Lemma 2.6 to (4.8), which then becomes

$$\sum_{k=1}^{\infty} s^k tr(A^k)/k + \sum_{k=1}^{\infty} t^k tr(B^k)/k = \sum_{k=1}^{\infty} tr[(sA+tB)^k]/k . \tag{4.9}$$

Equating the coefficients of s^2t^2 from both sides of (4.9) gives

$$0 = 4tr(A^2B^2) + 2tr(AB)^2 . \tag{4.10}$$

Since both A and B are symmetric (4.10) may be written as

$$2tr(AB)(AB)' + tr(AB+BA)(AB+BA)' = 0 . \tag{4.11}$$

As both $(AB)(AB)'$ and $(AB+BA)(AB+BA)'$ are positive semidefinite, (4.11) implies $AB = 0$, as required.

Proof of Theorem 4.1: Let $Q = \underset{\sim}{x}'A\underset{\sim}{x}$, $R = \underset{\sim}{x}'B\underset{\sim}{x}$ and let $\phi_1(s)$, $\phi_2(t)$ be the moment generating functions of Q and R respectively and let $\phi(s,t)$ be the joint moment generating function of Q and R. Then, since their moment generating functions exist, Q and R will be independent if and only if for all s and t sufficiently small,

$$\phi_1(s)\phi_2(t) = \phi(s,t) . \tag{4.12}$$

From (2.5), we obtain for all s and t sufficiently small so that the characteristic roots of $2sA\Sigma$, $2tB\Sigma$ and $2sA\Sigma + 2tB\Sigma$ are all less than one in absolute value,

$$\phi_1(s) = \frac{e^{s\underset{\sim}{\mu}'(I-2sA\Sigma)^{-1}A\underset{\sim}{\mu}}}{|I-2sA\Sigma|^{1/2}} \tag{4.13}$$

$$\phi_2(t) = \frac{e^{t\underset{\sim}{\mu}'(I-2tB\Sigma)^{-1}B\underset{\sim}{\mu}}}{|I-2tB\Sigma|^{1/2}} \tag{4.14}$$

$$\phi(s,t) = \frac{e^{\underset{\sim}{\mu}'(I-2sA\Sigma-2tB\Sigma)^{-1}(sA+tB)\underset{\sim}{\mu}}}{|I-2sA\Sigma-2tB\Sigma|^{1/2}} . \tag{4.15}$$

Using an argument similar to that used in the proof of Theorem 3.1, (4.12) will hold if and only if

$$|I-2sA\Sigma|.|I-2tB\Sigma| = |I-2sA\Sigma-2tB\Sigma| \tag{4.16}$$

and

$$s\underset{\sim}{\mu}'(I-2sA\Sigma)^{-1}A\underset{\sim}{\mu} + t\underset{\sim}{\mu}'(I-2tB\Sigma)^{-1}B\underset{\sim}{\mu} = \underset{\sim}{\mu}'(I-2sA\Sigma-2tB\Sigma)^{-1}(sA+tB)\underset{\sim}{\mu} . \tag{4.17}$$

Lemma 4.1 does not apply directly to (4.16) since $A\Sigma$ and $B\Sigma$ are not usually symmetric. If, however, we write $\Sigma = TT'$, then (4.16) is equivalent to

$$|I-2sT'AT|.|I-2tT'BT| = |I-2sT'AT-2tT'BT| . \tag{4.18}$$

We may take T to be a real $p\times r$ matrix where $r = rk(\Sigma)$. Using Lemma 4.1, we find that (4.18) and (4.16) hold if and only if

$$T'ATT'BT = 0 \tag{4.19}$$

or equivalently,

$$\Sigma A\Sigma B\Sigma = 0 . \tag{4.20}$$

Since we assume that s and t are sufficiently small so that the characteristic roots of $2sA\Sigma$, $2tB\Sigma$ and $2sA\Sigma + 2tB\Sigma$ are all less than one in absolute value, Lemma 2.5 applied to both sides of (4.17) gives

$$\sum_{k=0}^{\infty} \boldsymbol{\mu}'[(2s)^k(A\Sigma)^kA+(2t)^k(B\Sigma)^kB]\boldsymbol{\mu} = \sum_{k=0}^{\infty} \boldsymbol{\mu}'[(2sA\Sigma+2tB\Sigma)^k](sA+tB)\boldsymbol{\mu} . \qquad (4.21)$$

As each side of (4.21) is the same power series expansion, the coefficients of $s^jt^{j'}$, $j = 0,1,2,\ldots;$ $j' = 0,1,2,\ldots$ are equal on both sides of (4.21). When $j = j' = 1$, we obtain $\boldsymbol{\mu}'A\Sigma B\boldsymbol{\mu} + \boldsymbol{\mu}'B\Sigma A\boldsymbol{\mu} = 0$, and as $\boldsymbol{\mu}'A\Sigma B\boldsymbol{\mu} = \boldsymbol{\mu}'B\Sigma A\boldsymbol{\mu}$ (since the transpose of a scalar equals the scalar) we get

$$\boldsymbol{\mu}'A\Sigma B\boldsymbol{\mu} = 0 . \qquad (4.22)$$

When $j = j' = 2$, using (4.20) in (4.21) gives

$$\boldsymbol{\mu}'B\Sigma A\Sigma A\Sigma B\boldsymbol{\mu} + \boldsymbol{\mu}'A\Sigma B\Sigma B\Sigma A\boldsymbol{\mu} = 0 . \qquad (4.23)$$

As both $B\Sigma A\Sigma A\Sigma B = B\Sigma AT(B\Sigma AT)'$ and $A\Sigma B\Sigma B\Sigma A = A\Sigma BT(A\Sigma BT)'$ are positive semidefinite, (4.23) holds if and only if

$$T'A\Sigma B\boldsymbol{\mu} = \mathbf{0} , \quad T'B\Sigma A\boldsymbol{\mu} = \mathbf{0} , \qquad (4.24)$$

or equivalently,

$$\Sigma A\Sigma B\boldsymbol{\mu} = \mathbf{0} , \quad \Sigma B\Sigma A\boldsymbol{\mu} = \mathbf{0} . \qquad (4.25)$$

It is clear then that (4.21) is equivalent to (4.22) and (4.25) as the coefficients of $s^jt^{j'}$, $j = 1,2,3,\ldots;$ $j' = 1,2,3,\ldots$ on the right-hand side of (4.21) are all zero because of either (4.22) or (4.25). Also the coefficients of s^j, $j = 1,2,3,\ldots;$ and of $t^{j'}$, $j' = 1,2,3,\ldots$ are clearly equal on both sides of (4.21). Thus the necessary and sufficient conditions for the independence of the two quadratic forms $\mathbf{x}'A\mathbf{x}$ and $\mathbf{x}'B\mathbf{x}$ are (4.20), (4.22) and (4.25) which proves Theorem 4.1.

COROLLARY 4.1: (i) *If* $|\Sigma| \neq 0$, *condition* (4.4a) *is equivalent to*

$$A\Sigma B = 0 \ . \tag{4.26}$$

(ii) *If* A *is positive semidefinite, condition* (4.4a) *is equivalent to*

$$A\Sigma B(\Sigma, \mu) = 0 \ . \tag{4.27}$$

(iii) *If both* A *and* B *are positive semidefinite, condition* (4.4a) *reduces to* (4.26).

(iv) *If* $\Sigma = \sigma^2 I$, *where* σ^2 *is a scalar, then condition* (4.4a) *simplifies to*

$$AB = 0 \ . \tag{4.28}$$

<u>Proof</u>: The proof of (i) is immediate. To prove (ii), note that (4.27) implies (4.4a). Also (4.1) implies $\Sigma B\Sigma A\Sigma B\Sigma = 0$ which, since A is positive semidefinite, implies $A\Sigma B\Sigma = 0$ and (4.3) implies $\mu' B\Sigma A\Sigma B\mu = 0$ which implies $A\Sigma B\mu = 0$. Thus (ii) is proven. For (iii), note that (4.26) is sufficient for (4.4a). To see that it is also necessary, (4.1) implies $A\Sigma B\Sigma = 0$ from (ii), and this implies $A\Sigma B\Sigma A = 0$ and since B is positive semidefinite, $A\Sigma B = 0$. Finally, (iv) follows immediately.

Theorem 4.1 on the independence of quadratic forms is sometimes called Craig's Theorem or the Craig-Sakamoto Theorem. Craig (1943) claimed that if $x \sim N_p(0, I)$, then $x'Ax$ and $x'Bx$ are independent if and only if $AB = 0$. Sakamoto (1944b) claimed that if $x \sim N_p(0, \Sigma)$, Σ positive definite, then $x'Ax$ and $x'Bx$ would be independently distributed if and only if $A\Sigma B = 0$. Hotelling (1944) showed that Craig's proof was incorrect and attempted to give a correct proof, but

this also contained inaccuracies. Ogawa (1946a) also attempted to prove Craig's Theorem but apparently was unsatisfied as later, Ogawa (1949) states:

> "A.T. Craig (1943) and H. Sakamoto (1944b) showed that the ... condition $|I-sA-tB| = |I-sA|.|I-tB|$ is equivalent to $AB = 0$, but their proofs were incorrect. H. Hotelling (1944) also tried to prove this fact, but his proof was also not satisfactory. J. Ogawa (1946a) tried to derive these results ... but his proofs were also not satisfactory".

In his (1949) paper, Ogawa then proves Theorem 4.1 for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\Sigma)$, Σ positive definite.

Matusita (1949), who claims to have had the proof in 1943, and Aitken (1950) also proved Craig's Theorem for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\Sigma)$, Σ positive definite. Carpenter (1950) extended this result to noncentral variables; that is, if $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ positive definite, then a necessary and sufficient condition for $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ to be independent is still $A\Sigma B = 0$. This result was also proven by Ogawa (1950).

The result in the most general case when $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, Σ positive semidefinite as in Theorem 4.1, was first stated and proven by Ogasawara and Takahashi (1951). The proof we have given follows theirs to a great extent. Since 1951, there have been others who have restated and reproved this theorem. See, for example, Khatri (1963), Good (1963), and Rao and Mitra (1971).

Prior to Craig's Theorem, the criterion used to judge independence was effectively the factorization of the joint characteristic function. Cochran (1934) first stated this for the case when $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},I)$; i.e., $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are independent if and only if $|I-2isA-2itB| = |I-2isA|.|I-2itB|$ for all real s and t. Craig (1938) extended this result to the independence of q quadratic forms when $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\sigma^2 I)$. Aitken (1940) extended this result by proving that if $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\Sigma)$, Σ positive definite, a necessary and sufficient condition for $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ to be independent is $|I-s\Sigma A-t\Sigma B| = |I-s\Sigma A|\cdot|I-t\Sigma B|$, for all real s and t. Lancaster (1954), using cumulant generating functions, showed that when $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},I)$, a necessary and sufficient condition for the independence of $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ is $tr(sA)^j + tr(tB)^j = tr(sA+tB)^j$ for $j = 1,2,\ldots$; for all real s and t.

We extend Theorem 4.1 on the independence of two quadratic forms to Theorem 4.2 on the independence of two nonhomogeneous quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x} + 2\underset{\sim}{a}'\underset{\sim}{x} + a_1$ and $\underset{\sim}{x}'B\underset{\sim}{x} + 2\underset{\sim}{b}'\underset{\sim}{x} + b_1$. As the values of the constants a_1 and b_1 do not affect the independence of the nonhomogeneous forms, we do not lose generality by setting $a_1 = b_1 = 0$.

THEOREM 4.2: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$, *where* Σ *is positive semidefinite. Then a set of necessary and sufficient conditions for the independence of* $\underset{\sim}{x}'A\underset{\sim}{x} + 2\underset{\sim}{a}'\underset{\sim}{x}$ *and* $\underset{\sim}{x}'B\underset{\sim}{x} + 2\underset{\sim}{b}'\underset{\sim}{x}$ *is:*

$$\Sigma A\Sigma B\Sigma = 0 \tag{4.29}$$

$$\Sigma A\Sigma(B\underset{\sim}{\mu}+\underset{\sim}{b}) = \underset{\sim}{0}, \quad \Sigma B\Sigma(A\underset{\sim}{\mu}+\underset{\sim}{a}) = \underset{\sim}{0}\ , \tag{4.30}$$

$$(B\underset{\sim}{\mu}+\underset{\sim}{b})'\Sigma(A\underset{\sim}{\mu}+\underset{\sim}{a}) = 0\ . \tag{4.31}$$

<u>Proof</u>: We construct the new variable $\underset{\sim}{y} = \begin{pmatrix}\underset{\sim}{x}\\ 1\end{pmatrix}$. Then $\underset{\sim}{y} \sim N_{p+1}(\underset{\sim}{\mu}_o,\Sigma_o)$, where $\underset{\sim}{\mu}_o = \begin{pmatrix}\underset{\sim}{\mu}\\ 1\end{pmatrix}$ and $\Sigma_o = \begin{pmatrix}\Sigma & \underset{\sim}{0}\\ \underset{\sim}{0}' & 0\end{pmatrix}$. If $A_o = \begin{pmatrix}A & \underset{\sim}{a}\\ \underset{\sim}{a}' & 0\end{pmatrix}$ and $B_o = \begin{pmatrix}B & \underset{\sim}{b}\\ \underset{\sim}{b}' & 0\end{pmatrix}$, then $\underset{\sim}{y}'A_o\underset{\sim}{y} = \underset{\sim}{x}'A\underset{\sim}{x} + 2\underset{\sim}{a}'\underset{\sim}{x}$ and $\underset{\sim}{y}'B_o\underset{\sim}{y} = \underset{\sim}{x}'B\underset{\sim}{x} + 2\underset{\sim}{b}'\underset{\sim}{x}$.

The necessary and sufficient conditions for the independence of the nonhomogeneous quadratic forms are identical to the conditions for the independence of the two quadratic forms $\underset{\sim}{y}'A_o\underset{\sim}{y}$ and $\underset{\sim}{y}'B_o\underset{\sim}{y}$. These conditions are from Theorem 4.1:

$$\Sigma_o A_o \Sigma_o B_o \Sigma_o = 0 \tag{4.32}$$

$$\Sigma_o A_o \Sigma_o B_o \underset{\sim}{\mu}_o = \underset{\sim}{0}, \quad \Sigma_o B_o \Sigma_o A_o \underset{\sim}{\mu}_o = \underset{\sim}{0} \tag{4.33}$$

$$\underset{\sim}{\mu}_o' A_o \Sigma_o B_o \underset{\sim}{\mu}_o = 0 . \tag{4.34}$$

By substituting for Σ_o, A_o, B_o, $\underset{\sim}{\mu}_o$ in terms of Σ, A, B, $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{\mu}$ and simplifying, conditions (4.32), (4.33) and (4.34) become (4.29), (4.30) and (4.31) respectively.

COROLLARY 4.2: (i) *If* $\underset{\sim}{\mu} = \underset{\sim}{0}$, *the necessary and sufficient conditions in Theorem 4.2 reduce to*

$$\Sigma A \Sigma B \Sigma = 0, \quad \Sigma A \Sigma \underset{\sim}{b} = \underset{\sim}{0}, \quad \Sigma B \Sigma \underset{\sim}{a} = \underset{\sim}{0}, \quad \underset{\sim}{b}' \Sigma \underset{\sim}{a} = 0 . \tag{4.35}$$

(ii) *If* Σ *is positive semidefinite, the conditions reduce to*

$$A \Sigma B = 0, \quad A \Sigma \underset{\sim}{b} = \underset{\sim}{0}, \quad B \Sigma \underset{\sim}{a} = \underset{\sim}{0}, \quad \underset{\sim}{b}' \Sigma \underset{\sim}{a} = 0 . \tag{4.36}$$

(iii) *If* A *is positive semidefinite, the conditions simplify to*

$$A \Sigma B \Sigma = 0, \quad A \Sigma (B\underset{\sim}{\mu} + \underset{\sim}{b}) = \underset{\sim}{0}, \quad \Sigma B \Sigma \underset{\sim}{a} = \underset{\sim}{0}, \quad (B\underset{\sim}{\mu} + \underset{\sim}{b})' \Sigma \underset{\sim}{a} = 0 . \tag{4.37}$$

(iv) *If both* A *and* B *are positive semidefinite, the necessary and sufficient conditions for independence are* (4.36).

<u>Proof</u>: (i) Substituting $\underset{\sim}{\mu} = \underset{\sim}{0}$ into (4.29) - (4.31) gives (4.35).

(ii) Since Σ is positive definite, Σ^{-1} exists. Pre- and postmultiplying (4.29) - (4.31) by Σ^{-1} to remove the leading and/or the trailing Σ whenever necessary, gives (4.36).

(iii) When A is positive semidefinite, A may be expressed as FF'. Thus (4.29), $\Sigma A \Sigma B \Sigma = 0$ implies $\Sigma B \Sigma A \Sigma B \Sigma = 0$ or $\Sigma B \Sigma FF' \Sigma B \Sigma = (\Sigma B \Sigma F)(\Sigma B \Sigma F)' = 0$. This last statement implies $\Sigma B \Sigma F = 0$ and so $\Sigma B \Sigma FF' = \Sigma B \Sigma A = 0$, the first part of (4.37). Also premultiplying $\Sigma A \Sigma (B\mu + b) = 0$ by $(B\mu+b)'$ gives $(B\mu+b)' \Sigma FF' \Sigma (B\mu+b) = 0$, and so $A\Sigma(B\mu+b) = 0$, the second part of (4.37). The remaining two conditions follow from (4.30) and (4.31).

(iv) When, in addition to A being positive semidefinite, B is also positive semidefinite, the conditions (4.37) reduce further. B may be represented as GG'; then $A\Sigma B\Sigma = 0$ implies $A\Sigma B\Sigma A = A\Sigma GG'\Sigma A = 0$ and $A\Sigma B = 0$. This and the remainder of (4.37) reduce to (4.36).

The following lemma which was first stated and proven by Good (1963) will be useful in the proofs of subsequent theorems and corollaries. The result, as it appears in Good's paper is that if a set of quadratic expressions in normal variables are pairwise independent then they are mutually independent. The proof we give is stated in terms of quadratic forms and is not less general than Good's, as any quadratic expression may be represented as a quadratic form.

LEMMA 4.2: *Let* $x \sim N_p(\mu, \Sigma)$, Σ *positive semidefinite, and let* $x'A_1x$, $x'A_2x, \ldots, x'A_nx$ *be* n *quadratic forms which are pairwise independent. Then the quadratic forms are mutually independent.*

Proof: From Theorem 4.1, the pairwise independence of quadratic forms implies

$$\Sigma A_i \Sigma A_j \Sigma = 0 \tag{4.38a}$$

$$\Sigma A_i \Sigma A_j \mu = 0 \tag{4.38b}$$

$$\mu' A_i \Sigma A_j \mu = 0; \quad i,j = 1,2,\ldots,n; \; i \neq j . \tag{4.38c}$$

To show the mutual independence of the quadratic forms, the joint moment generating function $\phi(t_1,t_2,\ldots,t_n)$ must equal the product of the individual moment generating functions $\phi_1(t_1), \phi_2(t_2),\ldots, \phi_n(t_n)$, where from Chapter 2

$$\phi(t_1,t_2,\ldots,t_n) = \frac{e^{\mu'(I-\sum_{j=1}^{n} 2t_jA_j\Sigma)^{-1}\sum_{k=1}^{n} t_kA_k\mu}}{|I-\sum_{j=1}^{n} 2t_jA_j\Sigma|^{1/2}} \tag{4.39}$$

and

$$\phi_j(t_j) = \frac{e^{t_j\mu'(I-2t_jA_j\Sigma)^{-1}A_j\mu}}{|I-2t_jA_j\Sigma|^{1/2}} , \quad j=1,2,\ldots,n .$$

Thus, we must show that (4.39) equals

$$\phi_1(t_1)\phi_2(t_2)\ldots\phi_n(t_n) = \frac{e^{\sum_{j=1}^{n} t_j\mu'(I-2t_jA_j\Sigma)^{-1}A_j\mu}}{|I-2t_1A_1\Sigma|^{1/2}\ldots|I-2t_nA_n\Sigma|^{1/2}} ,$$

or

$$\mu'(I-\sum_{j=1}^{n} 2t_jA_j\Sigma)^{-1}\sum_{k=1}^{n} t_kA_k\mu = \sum_{j=1}^{n} t_j\mu'(I-2t_jA_j\Sigma)^{-1}A_j\mu \tag{4.40a}$$

and

$$|I-\sum_{j=1}^{n} 2t_jA_j\Sigma| = |I-2t_1A_1\Sigma|\ldots|I-2t_nA_n\Sigma| \ . \tag{4.40b}$$

For $t_1,t_2,\ldots,t_n$ sufficiently small so that the characteristic roots of $\sum_{j=1}^{n} 2t_jA_j\Sigma$ are all less than one in absolute value, Lemma 2.5 gives

$$(I-\sum_{j=1}^{n} 2t_jA_j\Sigma)^{-1} = \sum_{h=0}^{\infty} (\sum_{j=1}^{n} 2t_jA_j\Sigma)^h \ .$$

Therefore,

$$\mu'(I-\sum_{j=1}^{n} 2t_jA_j\Sigma)^{-1}\sum_{k=1}^{n} t_kA_k\mu = \sum_{h=0}^{\infty}\mu'(\sum_{j=1}^{n} 2t_jA_j\Sigma)^h(\sum_{k=1}^{n} t_kA_k\mu) \ . \tag{4.40c}$$

From (4.38a) - (4.38c) we see that terms with $j \neq k$ in the right-hand side of (4.40c) must be zero. Hence (4.40c) reduces to

$$\mu'(I-\sum_{j=1}^{n} 2t_jA_j\Sigma)^{-1}\sum_{k=1}^{n} t_kA_k\mu = \mu'[\sum_{j=1}^{n}\sum_{h=0}^{\infty}(2t_jA_j\Sigma)^h t_jA_j]\mu \ ,$$

which is the right-hand side of (4.40a) using Lemma 2.5.

Moreover,

$$|I-\sum_{j=1}^{n} 2t_jA_j\Sigma| = |I-2t_1A_1\Sigma-\sum_{j=2}^{n} 2t_jA_j\Sigma| = |I-2t_1T'A_1T-\sum_{j=2}^{n} 2t_jT'A_jT|, \tag{4.41a}$$

where $\Sigma = TT'$ and T is a $p\times r$ matrix of rank $r = rk(\Sigma)$. (4.38a) is then equivalent to $T'A_iTT'A_jT = 0;\ i,j = 1,\ldots,n,\ i \neq j$. In particular $T'A_1TT'A_jT = 0,\ j = 2,\ldots,n$ and

$$T'A_1T\ (\sum_{j=2}^{n} T'A_jT) = 0 \ . \tag{4.41b}$$

Using Lemma 4.1, (4.41b) implies that (4.41a) is equivalent to $|I-\sum_{j=1}^{n} 2t_j A_j \Sigma| = |I-2t_1 A_1 \Sigma| . |I-\sum_{j=2}^{n} 2t_j A_j \Sigma|$. Using the same argument recursively we obtain

$$|I-\sum_{j=1}^{n} 2t_j A_j \Sigma| = |I-2t_1 A_1 \Sigma| . |I-2t_2 A_2 \Sigma| \ldots |I-2t_n A_n \Sigma| ,$$

which is (4.40b) and thus Lemma 4.2 is proven.

The following Lemma 4.3, which we use in a subsequent corollary, is an immediate consequence of Lemma 4.2. It states that if a quadratic expression is independent of each member of a set of quadratic expressions then the quadratic expression is independent of the set.

LEMMA 4.3: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, Σ *positive semidefinite and let* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'B_1\underset{\sim}{x}$, $\underset{\sim}{x}'B_2\underset{\sim}{x}, \ldots, \underset{\sim}{x}'B_n\underset{\sim}{x}$ *be* $n + 1$ *quadratic forms such that* $\underset{\sim}{x}'A\underset{\sim}{x}$ *is independent of each* $\underset{\sim}{x}'B_i\underset{\sim}{x}$; $i = 1,2,\ldots,n$. *Then* $\underset{\sim}{x}'A\underset{\sim}{x}$ *is independent of the set of quadratic forms* $\{\underset{\sim}{x}'B_j\underset{\sim}{x};\ j=1,\ldots,n\}$.

Lemmas 4.2 and 4.3 may be extended to two sets of quadratic forms $\underset{\sim}{x}'A_1\underset{\sim}{x}$, $\underset{\sim}{x}'A_2\underset{\sim}{x}, \ldots, \underset{\sim}{x}'A_m\underset{\sim}{x}$ and $\underset{\sim}{x}'B_1\underset{\sim}{x}$, $\underset{\sim}{x}'B_2\underset{\sim}{x}, \ldots, \underset{\sim}{x}'B_n\underset{\sim}{x}$. If every $\underset{\sim}{x}'A_i\underset{\sim}{x}$ is independent of every $\underset{\sim}{x}'B_j\underset{\sim}{x}$ for $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$ then the set $\{\underset{\sim}{x}'A_i\underset{\sim}{x};\ i = 1,\ldots,m\}$ is independent of the set $\{\underset{\sim}{x}'B_j\underset{\sim}{x};\ j = 1,\ldots,n\}$.

COROLLARY 4.3: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, Σ *positive definite. Then:*

(i) *The necessary and sufficient condition for the linear combinations* $\underset{\sim}{a}'\underset{\sim}{x}$ *and* $\underset{\sim}{b}'\underset{\sim}{x}$ *to be independent is*

$$\underset{\sim}{a}'\Sigma\underset{\sim}{b} = 0 . \qquad (4.42a)$$

(ii) *The necessary and sufficient condition for the quadratic form* $\underset{\sim}{x}'A\underset{\sim}{x}$ *and the linear combination* $\underset{\sim}{b}'\underset{\sim}{x}$ *to be independent is*

$$A\Sigma\underset{\sim}{b} = \underset{\sim}{0} \; . \tag{4.42b}$$

(iii) *The necessary and sufficient condition for the quadratic form* $\underset{\sim}{x}'A\underset{\sim}{x}$ *and the set of* n *linear combinations* $B\underset{\sim}{x}$, *where* B *is an* $n \times p$ *matrix, to be independent is*

$$A\Sigma B' = 0 \; . \tag{4.42c}$$

<u>Proof</u>: (i) Substituting $A = 0$ and $B = 0$ in (4.29) - (4.31) gives $\underset{\sim}{a}'\Sigma\underset{\sim}{b} = 0$.

(ii) Substituting $B = 0$ and $\underset{\sim}{a} = \underset{\sim}{0}$ in (4.29) - (4.31) gives $A\Sigma\underset{\sim}{b} = \underset{\sim}{0}$.

(iii) The matrix B may be rewritten as

$$B = \begin{pmatrix} \underset{\sim}{b}_1' \\ \underset{\sim}{b}_2' \\ \vdots \\ \underset{\sim}{b}_n' \end{pmatrix}$$

and the vector $B\underset{\sim}{x}$ as

$$B\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{b}_1'\underset{\sim}{x} \\ \underset{\sim}{b}_2'\underset{\sim}{x} \\ \vdots \\ \underset{\sim}{b}_n'\underset{\sim}{x} \end{pmatrix} .$$

From Lemma 4.3, $\underset{\sim}{x}'A\underset{\sim}{x}$ is independent of $B\underset{\sim}{x}$ if it is independent of each $\underset{\sim}{b}_i'\underset{\sim}{x}$; $i = 1,\ldots,n$. From (ii) the necessary and sufficient condition for $\underset{\sim}{x}'A\underset{\sim}{x}$ to be independent of $\underset{\sim}{b}_i'\underset{\sim}{x}$ is $A\Sigma\underset{\sim}{b}_i = \underset{\sim}{0}$. As $\underset{\sim}{x}'A\underset{\sim}{x}$ and

$\underset{\sim}{b}_i'\underset{\sim}{x}$ must be independent for each $i = 1,2,\ldots,n,$

$$A\Sigma[\underset{\sim}{b}_1,\underset{\sim}{b}_2,\ldots,\underset{\sim}{b}_n] = [\underset{\sim}{0},\underset{\sim}{0},\ldots,\underset{\sim}{0}]$$

or

$$A\Sigma B' = 0 .$$

Theorem 4.2 and its corollaries appear widely throughout the literature, for example Aitken (1940, 1950), Kac (1945), Matérn (1949) Ogasawara and Takahashi (1951), Laha (1956), Khatri (1961a), Good (1963), Lukacs and Laha (1964), Styan (1970), Searle (1971) and Rao and Mitra (1971).

The first general treatment for quadratic forms in noncentral and possibly singular normal variables was by Ogasawara and Takahashi (1951). Good (1963) presents an alternative approach to obtain the necessary and sufficient conditions for the independence of the nonhomogeneous quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x} + \underset{\sim}{a}'\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x} + \underset{\sim}{b}'\underset{\sim}{x}$. He first establishes conditions for the independence of two quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$, of a quadratic form $\underset{\sim}{x}'A\underset{\sim}{x}$ and a linear combination $\underset{\sim}{b}'\underset{\sim}{x}$, and of two linear combinations $\underset{\sim}{a}'\underset{\sim}{x}$ and $\underset{\sim}{b}'\underset{\sim}{x}$. He then concludes that the necessary and sufficient conditions for the independence of the two nonhomogeneous forms are that $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'B\underset{\sim}{x}$; $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{b}'\underset{\sim}{x}$; $\underset{\sim}{a}'\underset{\sim}{x}$, $\underset{\sim}{x}'B\underset{\sim}{x}$; and $\underset{\sim}{a}'\underset{\sim}{x}$, $\underset{\sim}{b}'\underset{\sim}{x}$ must be pairwise independent.

We shall now obtain a set of necessary and sufficient conditions for any two bilinear forms $\underset{\sim}{x}'A\underset{\sim}{y}$ and $\underset{\sim}{x}'B\underset{\sim}{y}$ to be independent.

THEOREM 4.3: *Let* $\underset{\sim}{x},\underset{\sim}{y}$ *have a joint multivariate normal distribution such that* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}_1, \Sigma_{11})$ *and* $\underset{\sim}{y} \sim N_q(\underset{\sim}{\mu}_2, \Sigma_{22})$, *where both* Σ_{11} *and* Σ_{22} *are positive semidefinite. Also, let the cross-covariance matrix between* $\underset{\sim}{x}$ *and* $\underset{\sim}{y}$ *be* Σ_{12}. *Let* $\underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix}$ *and*

$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix}$. *Then a set of necessary and sufficient conditions for* $\underset{\sim}{x}'A\underset{\sim}{y}$ *and* $\underset{\sim}{x}'B\underset{\sim}{y}$, *where* A *and* B *are* p×q *matrices, to be independent is:*

$$\begin{pmatrix} \Sigma \\ \underset{\sim}{\mu}' \end{pmatrix} \begin{pmatrix} 0 & A \\ A' & 0 \end{pmatrix} \Sigma \begin{pmatrix} 0 & B \\ B' & 0 \end{pmatrix} (\Sigma, \underset{\sim}{\mu}) = 0 \ . \tag{4.43}$$

Proof: The random vector $\underset{\sim}{z} = \begin{pmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \end{pmatrix} \sim N_{p+q}(\underset{\sim}{\mu}, \Sigma)$.

Let $A_1 = (1/2)\begin{pmatrix} 0 & A \\ A' & 0 \end{pmatrix}$ and $B_1 = (1/2)\begin{pmatrix} 0 & B \\ B' & 0 \end{pmatrix}$. Then the set of necessary and sufficient conditions for $\underset{\sim}{x}'A\underset{\sim}{y}$ and $\underset{\sim}{x}'B\underset{\sim}{y}$ to be independent is equivalent to the set of necessary and sufficient conditions for $\underset{\sim}{z}'A_1\underset{\sim}{z}$ and $\underset{\sim}{z}'B_1\underset{\sim}{z}$ to be independent. From Theorem 4.1, (4.4a), $\underset{\sim}{z}'A_1\underset{\sim}{z}$ and $\underset{\sim}{z}'B_1\underset{\sim}{z}$ will be independent if and only if

$$\begin{pmatrix} \Sigma \\ \underset{\sim}{\mu}' \end{pmatrix} A_1 \Sigma B_1 (\Sigma, \underset{\sim}{\mu}) = 0 \ . \tag{4.44}$$

Substituting for A_1 and B_1 in terms of A and B gives (4.43) directly.

COROLLARY 4.4: (i) *If* $\underset{\sim}{\mu} = \underset{\sim}{0}$, *the necessary and sufficient condition for independence is*

$$\Sigma \begin{pmatrix} 0 & A \\ A' & 0 \end{pmatrix} \Sigma \begin{pmatrix} 0 & B \\ B' & 0 \end{pmatrix} \Sigma = 0 \ . \tag{4.45}$$

(ii) *If* Σ *is positive definite, the necessary and sufficient conditions for independence are:*

$$\begin{pmatrix} A\Sigma_{22}B' & A\Sigma_{12}'B \\ A'\Sigma_{12}B' & A'\Sigma_{11}B \end{pmatrix} = 0 \ . \tag{4.46}$$

(iii) *If* $\Sigma_{12} = 0$, *the necessary and sufficient conditions for independence are:*

$$\Sigma_{11}A\Sigma_{22}B'\Sigma_{11} = 0$$

$$\Sigma_{22}A'\Sigma_{11}B\Sigma_{22} = 0$$

$$\Sigma_{11}A\Sigma_{22}B'\mu_1 = 0 \qquad \Sigma_{11}B\Sigma_{22}A'\mu_1 = 0 \tag{4.47}$$

$$\Sigma_{22}A'\Sigma_{11}B\mu_2 = 0 \qquad \Sigma_{22}B'\Sigma_{11}A\mu_2 = 0$$

$$\mu_1'A\Sigma_{22}B'\mu_1 + \mu_2'A'\Sigma_{11}B\mu_2 = 0 \ .$$

Proof: (i) When $\mu = 0$, (4.43) clearly reduces to (4.45).

(ii) When Σ is positive definite, (4.44) reduces to $A_1\Sigma B_1 = 0$. Substituting for Σ, A_1, B_1 gives (4.46).

(iii) When $\Sigma_{12} = 0$; i.e., when x and y are uncorrelated, (4.43) reduces to (4.47).

Very little has appeared in the literature on the independence of bilinear forms. Part (ii) of Corollary 4.4 was proven by Aitken (1950). Craig (1947) proved the following result: let x, y be jointly multivariate normal such that $x \sim N_p(0, I)$, $y \sim N_p(0, I)$, and the covariance matrix between x and y is ρI. Then the necessary and sufficient for the independence of $\underset{\sim}{x}'A\underset{\sim}{y}$ and $\underset{\sim}{x}'B\underset{\sim}{y}$ is $AB = AB' = A'B = 0$. This result is readily obtained from (4.46).

An extension of the above is to consider the independence of two nonhomogeneous bilinear forms $\underset{\sim}{x}'A\underset{\sim}{y} + \underset{\sim}{a}_1'\underset{\sim}{x} + \underset{\sim}{a}_2'\underset{\sim}{y}$ and $\underset{\sim}{x}'B\underset{\sim}{y} + \underset{\sim}{b}_1'\underset{\sim}{x} + \underset{\sim}{b}_2'\underset{\sim}{y}$. We construct a new random variable $\underset{\sim}{z} = \begin{pmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \\ 1 \end{pmatrix}$ and matrices

$$A_1 = (1/2)\begin{pmatrix} 0 & A & \underset{\sim}{a}_1 \\ A' & 0 & \underset{\sim}{a}_2 \\ \underset{\sim}{a}_1' & \underset{\sim}{a}_2' & 0 \end{pmatrix}, \qquad B_1 = (1/2)\begin{pmatrix} 0 & B & \underset{\sim}{b}_1 \\ B' & 0 & \underset{\sim}{b}_2 \\ \underset{\sim}{b}_1' & \underset{\sim}{b}_2' & 0 \end{pmatrix}.$$

Then the necessary and sufficient conditions for the independence of the two nonhomogeneous bilinear forms are equivalent to the necessary and sufficient conditions for the independence of $\underset{\sim}{z}'A_1\underset{\sim}{z}$ and $\underset{\sim}{z}'B_1\underset{\sim}{z}$ which may be obtained in a similar manner to Theorem 4.3.

A different approach to obtain necessary and sufficient conditions for the independence of quadratic forms was taken by Matérn (1949), Kawada (1950), Lancaster (1954), Laha and Lukacs (1960) and Khatri (1961a).

Define the random variables x and y to be *uncorrelated of order* (r,s) *if*

$$\mathrm{cov}(x^i, y^i) = 0 \quad \text{for} \quad i = 1,\ldots,r; \quad j = 1,\ldots,s \; ,$$

where $\mathrm{cov}(x^i,y^j)$ is the covariance of x^i and y^i. Kawada (1950) showed that when $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},I)$, the necessary and sufficient condition for the quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ to be independent is that they must be uncorrelated of order (2,2). Moreover, if A is positive semidefinite, they must be uncorrelated of order (1,2). If both A and B are positive semidefinite then $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are uncorrelated [of order (1,1)].

Lancaster (1954) showed that when $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, I)$, $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are independent if and only if $\gamma_{ij} = 0$, for $i = 1,2$; $j = 1,2$, where γ_{ij} is the $(i,j)th$ cumulant of $(\underset{\sim}{x}'A\underset{\sim}{x}, \underset{\sim}{x}'B\underset{\sim}{x})$.

Laha and Lukacs (1960) showed that when $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, \Sigma)$, $\underset{\sim}{x}'A\underset{\sim}{x} + \underset{\sim}{a}'\underset{\sim}{x}$ and $\underset{\sim}{b}'\underset{\sim}{x}$ are independent if and only if they are uncorrelated of order (2,2).

Khatri (1961a) gives the most general results. He first shows the equivalence of Kawada's result with that of Lancaster's by proving that the quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are uncorrelated of order (r,s) if and only if $\gamma_{ij} = 0$ for $i = 1,\ldots,r$; $j = 1,\ldots,s$. He then proves that when $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, with Σ positive definite, $\underset{\sim}{x}'A\underset{\sim}{x} + \underset{\sim}{a}'\underset{\sim}{x}$ and $\underset{\sim}{b}'\underset{\sim}{x}$ are independent if and only if they are uncorrelated of order (2,2), or of order (1,2) if A is positive semidefinite. He also proves that if $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, \Sigma)$, Σ positive definite, then $\underset{\sim}{x}'A\underset{\sim}{x} + \underset{\sim}{a}'\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are independent if and only if they are uncorrelated of order (2,2) or of order (2,1) if B is positive semidefinite. Khatri points out that this last result "cannot be proved for noncentral normal variates" for the reason that "it is not possible to reduce in terms of finite higher order uncorrelation".

We conclude this chapter with miscellaneous results on the independence of quadratic forms that have appeared throughout the literature.

THEOREM 4.4: [Ogasawara and Takahashi (1951)] *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *with* Σ *positive definite. Suppose that* $\underset{\sim}{x}'A\underset{\sim}{x}$ *and* $\underset{\sim}{x}'B\underset{\sim}{x}$ *are independent and that* B *is positive semidefinite. Then for any positive semidefinite matrix* C *such that* B-C *is also positive semidefinite,* $\underset{\sim}{x}'A\underset{\sim}{x}$ *and* $\underset{\sim}{x}'C\underset{\sim}{x}$ *are independent.*

Proof: Since $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are independent and Σ is positive definite, part (i) of Corollary 4.1 gives $A\Sigma B = 0$. Since B-C is positive semidefinite so is $A\Sigma(B-C)\Sigma A$ or $A\Sigma B\Sigma A - A\Sigma C\Sigma A$. As $A\Sigma B = 0$, $-A\Sigma C\Sigma A$ must be positive semidefinite. But since C is positive semidefinite, so is $A\Sigma C\Sigma A$. Thus, $A\Sigma C\Sigma A = 0$ or, equivalently, $A\Sigma C = 0$. Therefore, by Corollary 4.1 (i), $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'C\underset{\sim}{x}$ are independent.

THEOREM 4.5: [Bhat (1962)] *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *with* Σ *positive definite or with* $\underset{\sim}{\mu} = \underset{\sim}{0}$ *and* Σ *positive semidefinite, and let* A *and* B *be positive semidefinite. Then* $\underset{\sim}{x}'C\underset{\sim}{x}$ *is distributed independently of* $\underset{\sim}{x}'A\underset{\sim}{x} + \underset{\sim}{x}'B\underset{\sim}{x}$ *if and only if it is distributed independently of the pair* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'B\underset{\sim}{x}$.

Bhat (1962) only considered Σ to be nonsingular; we show that Theorem 4.5 is also valid for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, \Sigma)$, where Σ is positive semidefinite. In both cases, it is obvious that the condition is sufficient.

Proof: When Σ is nonsingular, the independence of $\underset{\sim}{x}'C\underset{\sim}{x}$ and $\underset{\sim}{x}'(A+B)\underset{\sim}{x}$ implies, by Corollary 4.1(i), that $C\Sigma(A+B) = 0$. Postmultiplying by ΣC yields $C\Sigma(A+B)\Sigma C = 0$ or $C\Sigma A\Sigma C + C\Sigma B\Sigma C = 0$. As both matrices are positive semidefinite, they must both be equal to the zero matrix; i.e., $C\Sigma A\Sigma C = 0$ and $C\Sigma B\Sigma C = 0$. This last result gives $C\Sigma A = 0$ and $C\Sigma B = 0$ which by Corollary 4.1(i) implies that $\mathbf{x}'C\mathbf{x}$ is independent of $\mathbf{x}'A\mathbf{x}$ and of $\mathbf{x}'B\mathbf{x}$. When $\boldsymbol{\mu} = \mathbf{0}$ and Σ is positive semidefinite, the independence of $\mathbf{x}'C\mathbf{x}$ and $\mathbf{x}'(A+B)\mathbf{x}$ gives, by Corollary 4.4(ii), $\Sigma C\Sigma(A+B)\Sigma = 0$. Postmultiplying this by $C\Sigma$ yields $\Sigma C\Sigma(A+B)\Sigma C\Sigma = 0$. Using the same argument as above, we get $\Sigma C\Sigma A\Sigma C\Sigma = 0$ and $\Sigma C\Sigma B\Sigma C\Sigma = 0$ or $\Sigma C\Sigma A = 0$ and $\Sigma C\Sigma B = 0$. By Corollary 4.1(ii), $\mathbf{x}'C\mathbf{x}$ is independent of $\mathbf{x}'A\mathbf{x}$ and of $\mathbf{x}'B\mathbf{x}$. From Lemma 4.3, the pairwise independence of $\mathbf{x}'C\mathbf{x}$, $\mathbf{x}'A\mathbf{x}$ and $\mathbf{x}'C\mathbf{x}$, $\mathbf{x}'B\mathbf{x}$ implies that $\mathbf{x}'C\mathbf{x}$ is independent of the pair of quadratic forms $\mathbf{x}'A\mathbf{x}$, $\mathbf{x}'B\mathbf{x}$ (and hence of their sum).

Theorem 4.5 may be readily extended to the independence of a quadratic form and a set of n positive semidefinite quadratic forms. That is, if A_i; $i = 1,\ldots,n$ are positive semidefinite, then the quadratic form $\mathbf{x}'A\mathbf{x}$ is independent of the set of quadratic forms $\mathbf{x}'A_i\mathbf{x}$; $i = 1,\ldots,n$ if and only if it is independent of $\mathbf{x}'A_1\mathbf{x} + \mathbf{x}'A_2\mathbf{x} + \ldots + \mathbf{x}'A_n\mathbf{x}$. The proof follows by induction from Theorem 4.5 (or it may be proved in a direct manner, similar to the proof given for that theorem).

THEOREM 4.6: [Hotelling (1944), Laha (1956)] *Let* $\mathbf{x} \sim N_p(\mathbf{0}, I)$ *and let* $\mathbf{x}'A\mathbf{x}$ *and* $\mathbf{x}'B\mathbf{x}$ *be independent. Then there exists an orthogonal transformation* $\mathbf{y} = P\mathbf{x}$ *such that the resultant forms do not contain any variates in common.*

<u>Proof</u>: Since $\mathbf{x}'A\mathbf{x}$ and $\mathbf{x}'B\mathbf{x}$ are independent, $AB = BA = 0$. Hence there exists (cf. Mirsky, 1955, §10.6) an orthogonal matrix P such

that

$$PAP' = \Lambda \quad \text{and} \quad PBP' = \Delta ,$$

where Λ and Δ are diagonal. Hence $\Lambda\Delta = 0$ and so a nonzero diagonal element in Λ corresponds to a zero diagonal element in Δ and vice versa. Thus $\underset{\sim}{y}'PAP'\underset{\sim}{y}$ and $\underset{\sim}{y}'PBP'\underset{\sim}{y}$ do not contain any variates in common. Setting $\underset{\sim}{x} = P'\underset{\sim}{y}$ yields $\underset{\sim}{y} = P\underset{\sim}{x}$ and the proof is complete.

COROLLARY 4.5: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},\underset{\sim}{\Sigma})$, *with* $\underset{\sim}{\Sigma}$ *positive semidefinite of rank* $r \leq p$ *and let* $\underset{\sim}{x}'A\underset{\sim}{x}$ *and* $\underset{\sim}{x}'B\underset{\sim}{x}$ *be independent. Then there exists a random vector* $\underset{\sim}{y} \sim N_r(\underset{\sim}{0},I)$ *such that* $\underset{\sim}{x} = F\underset{\sim}{y}$, *where* F *is a* $p\times r$ *matrix, and such that the two quadratic forms will be functions of disjoint components of* $\underset{\sim}{y}$.

<u>Proof</u>: From Lemma 2.4, we know that there exists a $\underset{\sim}{z} \sim N_r(\underset{\sim}{0},I)$ such that $\underset{\sim}{x} = H\underset{\sim}{z}$ and H is a real $p\times r$ matrix. Then the quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'B\underset{\sim}{x}$ may be written as $\underset{\sim}{z}'H'AH\underset{\sim}{z}$ and $\underset{\sim}{z}'H'BH\underset{\sim}{z}$ respectively, and are still independent. Using Theorem 4.6, we know that there exists a vector $\underset{\sim}{y} = P\underset{\sim}{z}$ (or $\underset{\sim}{z} = P'\underset{\sim}{y}$), where P is orthogonal, which reduces the independent quadratic forms $\underset{\sim}{z}'H'AH\underset{\sim}{z}$ and $\underset{\sim}{z}'H'BH\underset{\sim}{z}$ to functions of disjoint components of $\underset{\sim}{y} = \{y_i\}$. As $\underset{\sim}{x} = H\underset{\sim}{z} = HP'\underset{\sim}{y}$, the independent quadratic forms $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'B\underset{\sim}{x}$ are functions of disjoint y_i and have no variates in common.

The next theorem, proved by Good (1963), generalizes some of the proceeding results. An outline of his proof is given.

THEOREM 4.7: *Let* $Q_1,Q_2,\ldots,Q_k$ *be quadratic expressions in*

$\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, \Sigma)$, *where* $rk(\Sigma) = r \leq p$. *If the* Q_i*'s are pairwise independent, then there exists a* $\underset{\sim}{z} \sim N_r(\underset{\sim}{0}, I)$ *for which* $\underset{\sim}{x}$ *is a linear transformation of* $\underset{\sim}{z}$ *and the* Q_i*'s are dependent on disjoint subsets of* $\underset{\sim}{z}$.

Proof: The proof is initially similiar to that of Theorem 4.6 in that we find a $\underset{\sim}{x} \sim N_r(\underset{\sim}{0}, I)$ such that $\underset{\sim}{x} = F\underset{\sim}{z}$. We then write $\underset{\sim}{z} = \underset{\sim}{z}^{(1)} + \underset{\sim}{z}^{(2)}$, where $\underset{\sim}{z}^{(1)}$ and $\underset{\sim}{z}^{(2)}$ are in orthogonal vector spaces dependent on the quadratic expressions. We then show that Q_1 is expressible in terms of the components of $\underset{\sim}{z}^{(1)}$, while $Q_2, \ldots, Q_k$ are functions of the components of $\underset{\sim}{x}^{(2)}$. This procedure is easily extended to all k quadratic expressions.

CHAPTER V

Cochran's Theorem

In Chapter III we obtained necessary and sufficient conditions for a quadratic form in normal variables to follow a chi-square distribution. In Chapter IV, we determined necessary and sufficient conditions for two quadratic forms to be independent. In this chapter we show how these two results interrelate. In particular, we derive a set of necessary and sufficient conditions for k quadratic forms to be pairwise independent and to individually follow chi-square distributions.

We also study, in detail, the situation, which occurs frequently in statistical analysis, where a quadratic form is expressed as the sum of a set of quadratic forms. In this case the conditions for having chi-square distributions and for being independent are sometimes equivalent. Indeed, any knowledge of the chi-squaredness of the quadratic forms may, on occasion, help determine their independence and vice versa. Also, a knowledge of relationships amongst the ranks of the quadratic forms is useful. These results are combined as *Cochran's Theorem*. We start with:

THEOREM 5.1: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *with* Σ *positive semidefinite and let* $\underset{\sim}{x}'A_i\underset{\sim}{x}$, $i=1,2,\ldots,k$, *be* k *quadratic forms. Then a set of necessary and sufficient conditions for the quadratic forms to be mutually independent while individually to follow chi-square distributions is*

$$\Sigma A_i \Sigma A_j \Sigma = \delta_{ij} \Sigma A_i \Sigma \tag{5.1}$$

$$\Sigma A_i \Sigma A_j \mu = \delta_{ij} \Sigma A_i \mu \qquad (5.2)$$

$$\mu' A_i \Sigma A_j \mu = \delta_{ij} \mu' A_i \mu , \qquad (5.3)$$

for $i,j = 1,2,\ldots,k$, *where* δ_{ij} *is the Kronecker delta* ($\delta_{ij} = 1$ *if* $i=j$ *and* $\delta_{ij} = 0$ *if* $i\neq j$).

We note that (5.1), (5.2) and (5.3) may be rewritten as

$$\begin{pmatrix} \Sigma \\ \mu' \end{pmatrix} A_i \Sigma A_j (\Sigma, \mu) = \delta_{ij} \begin{pmatrix} \Sigma \\ \mu' \end{pmatrix} A_i (\Sigma, \mu) . \qquad (5.4)$$

The proof is a straightforward application of Theorem 3.1 and Theorem 4.1. The conditions for chi-squaredness are those given when $i=j$ and $\delta_{ij} = 1$. The conditions for independence of $x'A_i x$, $x'A_j x$ are those when $i\neq j$ and $\delta_{ij} = 0$. Furthermore, the fact that the quadratic forms are pairwise independent implies, by Lemma 4.2, that the quadratic forms are mutually independent.

COROLLARY 5.1: (i) *If* $\mu = 0$, *the necessary and sufficient conditions reduce to* (5.1).

(ii) *If* Σ *is nonsingular the necessary and sufficient conditions reduce to* $A_i \Sigma A_j = \delta_{ij} A_i$.

(iii) *If* $\mu = 0$ *and* $\Sigma = I$, *the necessary and sufficient conditions reduce to* $A_i A_j = \delta_{ij} A_i$.

The proof of Corollary 5.1 follows by substituting the given conditions in (5.1), (5.2) and (5.3). Part (iii) was proven by Craig (1943).

COROLLARY 5.2: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$ *and let* $\underset{\sim}{x}'A_i\underset{\sim}{x} + 2\underset{\sim}{b}_i'\underset{\sim}{x} + c_i$; $i=1,\ldots,k,$ *be* k *nonhomogeneous quadratic forms. Then a set of necessary and sufficient conditions for them to be mutually independent and individually to follow chi-square distributions is*

$$\Sigma A_i \Sigma A_j \Sigma = \delta_{ij}\Sigma A_i \Sigma \tag{5.5}$$

$$\Sigma A_i \Sigma (A_j\underset{\sim}{\mu}+\underset{\sim}{b}_j) = \delta_{ij}\Sigma(A_j\underset{\sim}{\mu}+\underset{\sim}{b}_j) \tag{5.6}$$

$$(A_i\underset{\sim}{\mu}+\underset{\sim}{b}_i)'\Sigma(A_j\underset{\sim}{\mu}+\underset{\sim}{b}_j) = \delta_{ij}(\underset{\sim}{\mu}'A_i\underset{\sim}{\mu}+2\underset{\sim}{b}_i'\underset{\sim}{\mu}+c_i)\,, \tag{5.7}$$

for $i,j = 1,2,\ldots,k,$ *where* δ_{ij} *is the Kronecker delta.*

The proof of Corollary 5.2 follows directly from Theorem 5.1 by constructing a new variable $\underset{\sim}{z} = \begin{pmatrix}\underset{\sim}{x}\\ 1\end{pmatrix}$, and new matrices $B_i = \begin{pmatrix}A_i & \underset{\sim}{b}_i\\ \underset{\sim}{b}_i' & c_i\end{pmatrix}$ such that $\underset{\sim}{z}'B_i\underset{\sim}{z} = \underset{\sim}{x}'A_i\underset{\sim}{x} + 2\underset{\sim}{b}_i'\underset{\sim}{x} + c_i$. Applying Theorem 5.1 to the $\underset{\sim}{z}'B_i\underset{\sim}{z}$; $i=1,\ldots,k$ and resubstituting gives (5.5), (5.6) and (5.7). This result may also be obtained from Corollary 3.9 and Theorem 4.2.

Although both Theorem 5.1 and its two corollaries are straightforward, they do serve as a good starting point for Cochran's Theorem as they show the similarity in structure between conditions for chi-squaredness and conditions for independence.

Cochran's Theorem as given by Cochran (1934) is the following. Let $\underset{\sim}{x} \sim N_p(\underset{\sim}{0},I)$ and let A_i, $i=1,\ldots,k$ be symmetric matrices such that $\sum_{i=1}^{k} A_i = I$ and let $\mathrm{rk}(A_i) = r_i$. Then $\underset{\sim}{x}'A_i\underset{\sim}{x}$ follows a chi-square distribution with r_i degrees of freedom and is independent of the other $\underset{\sim}{x}'A_j\underset{\sim}{x}$ if and only if $\sum_{i=1}^{k} r_i = p$. Note how much simpler this condition is to state (and sometimes to verify) than $A_i^2 = A_i$ and $A_iA_j = 0$ for $i,j = 1,\ldots,k$ and $i \neq j$. We shall prove a more general version of this theorem as given by Styan (1970).

THEOREM 5.2: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, *with* Σ *positive semidefinite, and let* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'A_i\underset{\sim}{x}$, $i=1,\ldots,k$ *be* $k+1$ *quadratic forms such that* $\underset{\sim}{x}'A\underset{\sim}{x} = \sum_{i=1}^{k} \underset{\sim}{x}'A_i\underset{\sim}{x}$. *Also, let* $r = rk(\Sigma A \Sigma)$ *and* $r_i = rk(\Sigma A_i \Sigma)$, $i=1,\ldots,k$. *If* (I) Σ *is nonsingular or* (II) Σ *is singular and* $\underset{\sim}{\mu} = \underset{\sim}{0}$ *or* (III) Σ *is singular,* $\underset{\sim}{\mu}$ *not necessarily* $\underset{\sim}{0}$, *and the* A_i *are positive semidefinite then*

$$\text{(a) and (d) imply (b) and (c)} \tag{5.8}$$

$$\text{(a) and (b) imply (c) and (d)} \tag{5.9}$$

$$\text{(a) and (c) imply (b) and (d)} \tag{5.10}$$

$$\text{(b) and (c) imply (a) and (d)}\,, \tag{5.11}$$

where

(a) $\underset{\sim}{x}'A\underset{\sim}{x} \sim \chi_r^2(\delta^2)$

(b) $\underset{\sim}{x}'A_i\underset{\sim}{x} \sim \chi_{r_i}^2(\delta_i^2)$; $i = 1,\ldots,k$

(c) $\underset{\sim}{x}'A_1\underset{\sim}{x}, \underset{\sim}{x}'A_2\underset{\sim}{x}, \ldots, \underset{\sim}{x}'A_k\underset{\sim}{x}$ *mutually independent*

(d) $r = \sum_{i=1}^{k} r_i$.

If (IV) Σ *is singular,* $\underset{\sim}{\mu} \neq \underset{\sim}{0}$, *and the* A_i *are not all positive semidefinite then* (5.9) *and* (5.11) *still hold, while* (5.8) *and* (5.10) *need not hold.*

<u>Proof</u>: Let $\Sigma = TT'$, where T is a $p\times m$ matrix and $m = rk(\Sigma)$. When Σ is nonsingular so is T and then $r = rk(A)$ and $r_i = rk(A_i)$.

Under (I) and (II) we may rewrite (a), (b) and (c) as

$$\text{(a')} \quad T'AT = (T'AT)^2 ,$$

$$\text{(b')} \quad T'A_iT = (T'A_iT)^2 ; \quad i = 1,\dots,k ,$$

$$\text{(c')} \quad T'A_iTT'A_jT = 0 ; \quad i,j = 1,\dots,k ; \quad i \neq j ,$$

using Theorem 3.1 and Theorem 4.1. Under (III) we need in addition to (a') and (b')

$$\text{(a'')} \quad \mu'A\Sigma A\mu = \mu'A\mu$$

$$\text{(b'')} \quad \mu'A_i\Sigma A_i\mu = \mu'A_i\mu, \quad i = 1,\dots,k .$$

Note also that $r_i = rk(\Sigma A_i \Sigma) = rk(T'A_iT)$. We first prove (5.8) - (5.11) with (a), (b), (c) replaced by (a'), (b') and (c'). We then prove that when these hold (a") and (b") are equivalent. We may write (cf. Styan, 1970) $T'A_iT = U_iV_i'$, where both U_i and V_i are real $m\times r_i$ matrices and have full column rank for $i=1,\dots,k$. Then

$$T'AT = \sum_{i=1}^{k} T'A_iT = \sum_{i=1}^{k} U_iV_i' = UV' , \qquad (5.12)$$

where $U = (U_1,U_2,\dots,U_k)$ and $V = (V_1,V_2,\dots,V_k)$ are both $m\times r$. From (a') $UV'UV' = UV'$. As from (d) U and V have full column rank, Lemma 3.2 gives $V'U = I$. As

$$V'U = \begin{pmatrix} V_1' \\ V_2' \\ \vdots \\ V_k' \end{pmatrix} (U_1,\dots,U_k) = \begin{pmatrix} V_1'U_1 & \dots & V_1'U_k \\ V_2'U_1 & \dots & V_2'U_k \\ \vdots & & \vdots \\ V_k'U_1 & \dots & V_k'U_k \end{pmatrix} , \qquad (5.13)$$

we get $V_i'U_i = I_{r_i}$, for $i=1,\dots,k$, and $V_i'U_j = 0$, for $i,j = 1,\dots,k$; $i \neq j$. Thus $U_iV_i'U_iV_i' = U_iV_i'$ and (b') and (c') follow. To show (5.9) we use the result that the rank of a symmetric idempotent matrix is equal to its trace. (a') and (b') imply $r = rk(T'AT) = tr(T'AT) = tr(\sum_{i=1}^{k} T'A_iT) = \sum_{i=1}^{k} tr(T'A_iT) = \sum_{i=1}^{k} rk(T'A_iT) = \sum_{i=1}^{k} r_i$ and thus (d) holds. By (5.8) (c') also holds. When (a') and (c') hold,

$$T'AT = \sum_{i=1}^{k} T'A_iT = (\sum_{i=1}^{k} T'A_iT)^2 = \sum_{i=1}^{k} (T'A_iT)^2 , \text{ or}$$

$$\sum_{i=1}^{k} T'A_iT = \sum_{i=1}^{k} (T'A_iT)^2 . \qquad (5.14)$$

Multiplying (5.14) by $T'A_jT$, for fixed j, gives

$$(T'A_jT)^2 = (T'A_jT)^3 . \qquad (5.15)$$

Since the rank of a symmetric matrix equals the rank of any power of it, Lemma 3.2 applied to (5.15) gives

$$(T'A_jT)^2 = T'A_jT , \quad j = 1,\dots,k . \qquad (5.16)$$

Thus (b') holds and hence (d). When (b') and (c') hold

$$(T'AT)^2 = (\sum_{i=1}^{k} T'A_iT)^2 = \sum_{i=1}^{k} (T'A_iT)^2 + \sum_{i \neq j} T'A_iTT'A_jT$$

$$= \sum_{i=1}^{k} (T'A_iT)^2 = \sum_{i=1}^{k} T'A_iT = T'AT . \qquad (5.17)$$

Thus, (a') holds and hence (d) also holds. When (a'), (b'), (c') and (d) hold, then clearly (b") implies (a"). When each A_i is positive semidefinite, we may write $A_i = B_i B_i'$ for $i=1,\ldots,k$ and then (b') gives $T'B_iB_i'TT'B_jB_j'T = 0$ for $i \neq j$; therefore

$$\mathrm{tr}(T'B_iB_i'TT'B_jB_j'T) = 0 \ . \tag{5.18}$$

Also,

$$\begin{aligned} \mathrm{tr}(T'B_iB_i'TT'B_jB_j'T) &= \mathrm{tr}(B_j'TT'B_iB_i'TT'B_j) \\ &= \mathrm{tr}(B_j'TT'B_i)(B_j'TT'B_i)' \end{aligned} \tag{5.19}$$

and $B_jTT'B_i = 0$ for $i,j = 1,\ldots,k$; $i \neq j$. Thus

$$\begin{aligned} \mu'A\Sigma A\mu &= \mu' \sum_{i=1}^{k} A_i \Sigma \sum_{j=1}^{k} A_j \mu \\ &= \sum_{i=1}^{k} \sum_{j=1}^{k} \mu' B_iB_i'TT'B_jB_j'\mu \\ &= \sum_{i=1}^{k} \mu' B_iB_i'TT'B_iB_i'\mu \ . \end{aligned} \tag{5.20}$$

From (a"), $\mu'A\Sigma A\mu = \mu'A\mu$, and (5.20), we get

$$\sum_{i=1}^{k} \mu'B_iB_i'TT'B_iB_i'\mu = \sum_{i=1}^{k} \mu'B_iB_i'\mu$$

or equivalently,

$$\sum_{i=1}^{k} \mu'B_i(I-B_i'TT'B_i)B_i\mu = 0 \ . \tag{5.21}$$

As $\mathrm{ch}(I-B_i'TT'B_i) = 1 - \mathrm{ch}(B_i'TT'B_i) = 1 - \mathrm{ch}(T'A_iT)$ and as $\mathrm{ch}(T'A_iT)$ are 0 or 1, the characteristic roots of $I - B_i'TT'B_i$ are either 1 or 0. Thus $I - B_i'TT'B_i$ is positive semidefinite for all i and $\mu'B_i(I-B_i'TT'B_i)B_i'\mu \geq 0$. From the above and (5.21) we must conclude that $\mu'B_i(I-B_i'TT'B_i)B_i'\mu = 0$, or $\mu'A_i\Sigma A_i\mu = \mu'A_i\mu$ for all $i=1,\ldots,k$, as required.

Under (IV), for (5.9) we must show that (a'), (b'), (c') and (d) together with $\Sigma A\Sigma A\mu = \Sigma A\mu$ and $\Sigma A_i\Sigma A_i\mu = \Sigma A_i\mu$ imply $\Sigma A_i\Sigma A_j\mu = 0$ for $i\neq j$. Premultiplying $\Sigma A_i\Sigma A_i\mu = \Sigma A_i\mu$ by ΣA_j and using (c') gives $\Sigma A_j\Sigma A_i\mu = 0$ as required. For (5.11) we must show that $\Sigma A_i\Sigma A_i\mu = \Sigma A_i\mu$ and $\Sigma A_i\Sigma A_j\mu = 0$ imply $\Sigma A\Sigma A\mu = \Sigma A\mu$. $\Sigma A\Sigma A\mu = \sum_{i,j=1}^{k} \Sigma A_i\Sigma A_j\mu = \sum_{i=1}^{k} \Sigma A_i\Sigma A_i\mu = \sum_{i=1}^{k} \Sigma A_i\mu = \Sigma A\mu$ and (5.11) is proven. To show that (5.8) and (5.10) need not hold, consider this counter-example.

Let $\mu = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $A = \Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Then $A = A_1 + A_2$ and Σ (and A) may be written as TT' where $T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$. Note that Σ, A and A_1 are idempotent and that $T'T = I_2$, so that $T'ATT'AT = T'AT$, $\mu'A\Sigma A\Sigma = \mu'A\Sigma$ and $\mu'A\Sigma A\mu = \mu'A\mu = 0$. Thus (a) holds. Also $A_1\Sigma A_2 = A_2\Sigma A_1 = 0$, so that (c) holds. Finally $rk(T'AT) = 2 = rk(T'A_1T) + rk(T'A_2T)$ and (d) holds. However, $\mu'A_1\Sigma A_1\mu = 0 \neq \mu'A_1\mu = 1$ so that (b) does not hold. Therefore (5.8) and (5.10) are shown not to hold in general under (IV). Thus the proof of Theorem 5.2 is completed.

The result as stated in Theorem 5.2 has been given and proved by Styan (1970), Searle (1971) and Rao and Mitra (1971). Ogasawara and

Takahashi (1951) have also established most of the above results.

Madow (1940) extended Cochran's result for $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, I)$. Sakamoto (1944b), Ogawa (1946b) and Matérn (1949) considered Cochran's Theorem for $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, \Sigma)$, with Σ nonsingular. Graybill and Marsaglia (1957) and Chipman and Rao (1964) extended the results to $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, Σ nonsingular. In addition, Craig (1938), Aitken (1950), Dieulefait (1951), Nelder (1951), Lancaster (1954), Banerjee (1964), Rao (1965), Loynes (1966) and Luther (1965), have all proven Cochran's initial result as well as other related results.

It should be noted that many of these results correspond to results in matrix algebra. In fact, Theorem 5.2 may be so restated. Although credit for this theorem has always been given to Cochran, it should be noted that Fisher (1925) gave a result which foreruns Cochran's Theorem. In fact, Rao (1965) refers to Cochran's result as the Fisher-Cochran Theorem. Fisher's result is as follows. Let $\underset{\sim}{x} \sim N_p(\underset{\sim}{0}, I)$. If $\underset{\sim}{z} = P\underset{\sim}{x}$ where P is $h \times p$ such that $PP' = I_h$, then $\underset{\sim}{x}'\underset{\sim}{x} - \underset{\sim}{z}'\underset{\sim}{z} \sim \chi^2_{p-h}$ independently of $\underset{\sim}{z}$.

We now state and prove a result given by Hogg and Craig (1958).

THEOREM 5.3: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ *with* (I) Σ *nonsingular or* (II) Σ *singular and* $\underset{\sim}{\mu} = \underset{\sim}{0}$, *and let* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'A_i\underset{\sim}{x}$; $i=1,\ldots,k$ *be* $k+1$ *quadratic forms such that* $A = \sum_{i=1}^{k} A_i$. *If* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'A_i\underset{\sim}{x}$; $i=1,\ldots,k-1$ *follow chi-square distributions with degrees of freedom* r, r_i; $i=1,\ldots, k-1$, *respectively, and if* A_k *is positive semidefinite then* $\underset{\sim}{x}'A_k\underset{\sim}{x}$ *also follows a chi-square distribution with* $r - \sum_{i=1}^{k-1} r_i$ *degrees of freedom and* $\underset{\sim}{x}'A_1\underset{\sim}{x}$, $\underset{\sim}{x}'A_2\underset{\sim}{x}$, $\ldots$, $\underset{\sim}{x}'A_k\underset{\sim}{x}$ *are independent.*

Proof: Write $\Sigma = TT'$ where T has rank $s = rk(\Sigma)$. Then under (I) or (II) the conditions that $x'Ax$, $x'A_ix$ follow chi-square distributions are

$$T'AT = (T'AT)^2 \tag{5.22}$$

$$T'A_iT = (T'A_iT)^2 \ ; \quad i=1,\ldots,k-1, \tag{5.23}$$

respectively, where $r = rk(T'AT)$, $r_i = rk(T'A_iT)$. We shall prove this result by induction on k. Let $k=2$; i.e., $A = A_1 + A_2$ and $x'Ax$, $x'A_1x$ follow chi-square distributions. Also, A_2 is positive semidefinite. We have

$$\begin{aligned} T'AT = (T'AT)^2 &= (T'A_1T+T'A_2T)^2 \\ &= (T'A_1T)^2 + T'A_1TT'A_2T + T'A_2TT'A_1T + (T'A_2T)^2. \end{aligned} \tag{5.24}$$

Using (5.22) and (5.23) in (5.24) gives

$$T'A_1T + T'A_2T = T'A_1T + T'A_1TT'A_2T + T'A_2TT'A_1T + (T'A_2T)^2$$

or, equivalently,

$$T'A_2T = T'A_1TT'A_2T + T'A_2TT'A_1T + (T'A_2T)^2 \ . \tag{5.25}$$

Premultiplying (5.25) by $T'A_1T$ and using (5.23) to eliminate $T'A_1TT'A_2T$ from both sides of the equation gives

$$0 = T'A_1TT'A_2TT'A_1T + T'A_1T(T'A_2T)^2 \ . \tag{5.26}$$

Equating the traces of both sides of (5.26) gives

$$0 = tr(T'A_1TT'A_2TT'A_1T) + tr(T'A_2TT'A_1TT'A_2T) \ . \tag{5.27}$$

As both matrix expressions are positive semidefinite they must each equal the zero matrix and, in particular,

$$T'A_1TT'A_2T = 0 \ . \tag{5.28}$$

Substituting this into (5.25) gives

$$T'A_2T = (T'A_2T)^2 \tag{5.29}$$

and thus $\underset{\sim}{x}'A_2\underset{\sim}{x}$ follows a chi-square distribution. Using (5.9) in Theorem 5.2, we obtain $r_2 = r - r_1$ and that $\underset{\sim}{x}'A_1\underset{\sim}{x}$ and $\underset{\sim}{x}'A_2\underset{\sim}{x}$ are independent. Let us assume the theorem is true for $k = n$. Now let $k = n + 1$; i.e., $A = \sum_{i=1}^{n+1} A_i$, where $\underset{\sim}{x}'A\underset{\sim}{x}$ and $\underset{\sim}{x}'A_i\underset{\sim}{x}$; $i=1,\ldots,n$ follow chi-square distributions and A_{n+1} is positive semidefinite. Let $B = A_n + A_{n+1}$. Then under (I) B is also positive semidefinite and $A = \sum_{i=1}^{n-1} A_i + B$. From our induction assumption for $k=n$, $\underset{\sim}{x}'B\underset{\sim}{x}$ has a chi-square distribution. But $B = A_n + A_{n+1}$, where $\underset{\sim}{x}'A_n\underset{\sim}{x}$ follows a chi-square distribution and A_{n+1} is positive semidefinite. Therefore $\underset{\sim}{x}'A_{n+1}\underset{\sim}{x}$ must also follow a chi-square distribution as this is the case where $k=2$. Under (II), from Lemma 2.4 there exists a $\underset{\sim}{y} \sim N_s(\underset{\sim}{0}, I)$ such that $\underset{\sim}{x} = T\underset{\sim}{y}$. Then $\underset{\sim}{y}'T'AT\underset{\sim}{y} = \sum_{i=1}^{n+1} \underset{\sim}{y}'T'A_iT\underset{\sim}{y}$, where $\underset{\sim}{y}'T'AT\underset{\sim}{y}$, $\underset{\sim}{y}'T'A_iT\underset{\sim}{y}$, $i=1,\ldots,n$ have chi-square distribution and $T'A_{n+1}T$ is positive semidefinite. As $\underset{\sim}{y}$ has a nonsingular covariance matrix, the first part of this proof shows that $\underset{\sim}{y}'T'A_{n+1}T\underset{\sim}{y} = \underset{\sim}{x}'A_{n+1}\underset{\sim}{x}$ must follow a chi-square distribution. As under (I) or (II) $\underset{\sim}{x}'A_i\underset{\sim}{x}$ has a chi-square distribution for all i, they must again be independent and $r_{n+1} = r - \sum_{i=1}^{n} r_i$. Our proof by induction is therefore complete.

Part (I) of Theorem 5.3 appeared in Hogg and Craig (1958) and was later proven differently by Banerjee (1964). It was also obtained earlier by Ogasawara and Takahashi (1951). Note that this theorem makes (5.9) of Theorem 5.2 more general. We now conclude this chapter by extending Theorem 5.3 with the following:

COROLLARY 5.3: *Let* $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ *with* (I) Σ *nonsingular or* (II) Σ *singular and* $\underset{\sim}{\mu} = \underset{\sim}{0}$, *and let* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'A_i\underset{\sim}{x}$; $i=1,\ldots,k+\ell$ *be* $k+\ell+1$ *quadratic forms such that* $A = \sum_{i=1}^{k+\ell} A_i$. *Also let* $\underset{\sim}{x}'A\underset{\sim}{x}$, $\underset{\sim}{x}'A_i\underset{\sim}{x}$; $i=1,\ldots,k$ *follow chi-square distributions and let* A_j; $j = k+1,\ldots,k+\ell$ *be positive semidefinite. Let* $\Sigma = TT'$, *where* T *has full column rank. If either*

(a) $\mathrm{rk}(\sum_{j=k+1}^{k+\ell} T'A_jT) = \sum_{j=k+1}^{k+\ell} \mathrm{rk}(T'A_jT)$, *or*

(b) $\underset{\sim}{x}'A_i\underset{\sim}{x}$, $\underset{\sim}{x}'A_j\underset{\sim}{x}$ *are independent for* $i,j = k+1,\ldots,k+\ell$; $i \neq j$,

then each $\underset{\sim}{x}'A_j\underset{\sim}{x}$, $j = k+1,\ldots,k+\ell$ *follows a chi-square distribution and all the* $\underset{\sim}{x}'A_i\underset{\sim}{x}$, $i=1,\ldots,k+\ell$ *are independent.*

<u>Proof</u>: Under either (I) or (II), since each A_j, $j = k+1,\ldots,k+\ell$ is positive semidefinite, the sum $\sum_{j=k+1}^{k+\ell} A_j$ is also positive semidefinite. Thus Theorem 5.3 shows that $\underset{\sim}{x}' \sum_{j=k+1}^{k+\ell} A_j\underset{\sim}{x}$ follows a chi-square distribution. This and either (a) or (b) imply by (5.8) and (5.10) respectively, that $\underset{\sim}{x}'A_j\underset{\sim}{x}$ for $j = k+1,\ldots,k+\ell$ follow chi-square distributions. As each $\underset{\sim}{x}'A_i\underset{\sim}{x}$, $i=1,\ldots,k+\ell$ and $\underset{\sim}{x}'A\underset{\sim}{x}$ follow chi-square distributions, reapplying (5.9) completes the proof.

REFERENCES AND BIBLIOGRAPHY

Aitken, A.C. (1940). On the independence of linear and quadratic forms in samples of normally distributed variates. *Proceedings of the Royal Society of Edinburgh*, 60, 40-46. [Corrigendum: *Series A*, 62 (1948), 277]

Aitken, A.C. (1950). On the statistical independence of quadratic forms in normal variates. *Biometrika*, 37, 93-96.

Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, Inc., New York.

Anderson, T.W.; Das Gupta, S., & Styan, G.P.H. (1972). *A Bibliography of Multivariate Statistical Analysis*. Oliver & Boyd, Edinburgh.

Baldessari, B. (1967). The distribution of a quadratic form of normal random variables. *The Annals of Mathematical Statistics*, 38, 1700-1704.

Banerjee, K.S. (1964). A note on idempotent matrices. *The Annals of Mathematical Statistics*, 35, 880-882.

Bhapkar, V.P. (1970). The invariance and distribution of the quadratic form $\underset{\sim}{x}'\underset{\sim}{\Sigma}^{-}\underset{\sim}{x}$. Technical Report No. 4, Department of Statistics, University of Kentucky, 8pp.

Bhat, B.R. (1959). On the distribution of various sums of squares in an analysis of variance table for different classifications with correlated and non-homogeneous errors. *Journal of the Royal Statistical Society Series B*, 21, 114-119.

Bhat, B.R. (1962). On the distribution of certain quadratic forms in normal variates. *Journal of the Royal Statistical Society Series B*, 24, 148-151.

Carpenter, O. (1950). Note on the extension of Craig's theorem to non-central variates. *The Annals of Mathematical Statistics*, 21, 455-457.

Chipman, J.S. & Rao, M.M. (1964). Projections, generalized inverses and quadratic forms. *Journal of Mathematical Analysis and Applications*, 9, 1-11.

Cochran, W.G. (1934). The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. *Proceedings of the Cambridge Philosophical Society*, 30, 178-191.

Craig, A.T. (1936). Note on a certain bilinear form that occurs in statistics. *American Journal of Mathematics*, 58, 864-866.

Craig, A.T. (1936). A certain mean-value problem in statistics. *Bulletin of the American Mathematical Society*, 42, 670-674.

Craig, A.T. (1938). On the independence of certain estimates of variance. *The Annals of Mathematical Statistics*, 9, 48-55.

Craig, A.T. (1943). Note on the independence of certain quadratic forms. *The Annals of Mathematical Statistics*, 14, 195-197.

Craig, A.T. (1947). Bilinear forms in normally correlated variables. *The Annals of Mathematical Statistics*, 18, 565-573.

Dieulefait, E. (1951). Sobre las formas cuadraticas a variables aleatorias. (About quadratic forms with random variables). *Anales de la Sociedad Cientifica Argentina*, 151, 167-172.

Fisher, R.A. (1925). Applications of "Student's" distribution. *Metron*, 5, 90-104.

Goldberger, A.S. (1964). *Econometric Theory*. John Wiley & Sons, Inc., New York.

Good, I.J. (1963). On the independence of quadratic expressions (with an appendix by L.R. Welch). *Journal of the Royal Statistical Society Series B*, 25, 377-382. [Corrigenda: 28 (1966), 584]

Good, I.J. (1969). Conditions for a quadratic form to have a chi-squared distribution. *Biometrika*, 56, 215-216. [Correction: 57 (1970), 225]

Graybill, F.A. (1961). *An Introduction to Linear Statistical Models. Volume I*. McGraw-Hill Book Company, New York.

Graybill, F.A. & Marsaglia, G. (1957). Idempotent matrices and quadratic forms in the general linear hypothesis. *The Annals of Mathematical Statistics*, 28, 678-686.

Hogg, R.V. & Craig, A.T. (1958). On the decomposition of certain χ^2 variables. *The Annals of Mathematical Statistics*, 29, 608-610.

Hogg, R.V. & Craig, A.T. (1970). *Introduction to Mathematical Statistics*. Third Edition. Macmillan Company, New York.

Hotelling, H. (1944). Note on a matric theorem of A.T. Craig. *The Annals of Mathematical Statistics*, 15, 427-429.

Hultquist, R. (1966). Diagramming theorems relative to quadratic forms. *The American Statistician*, 20 (5), 31.

Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419-426.

James, G.S. (1952). Notes on a theorem of Cochran. *Proceedings of the Cambridge Philosophical Society*, 48, 443-446.

Kac, M. (1945). A remark on independence of linear and quadratic forms involving independent Gaussian variables. *The Annals of Mathematical Statistics*, 16, 400-401.

Kamalov, M.K. (1953). Condition for independence of ratios of quadratic forms. (In Russian) *Akademija Nauk Uzbekskoĭ SSR. Trudy Instituta Matematiki i Mehaniki im. V.I. Romanovskogo*, 11, 88-91.

Kamalov, M.K. (1955). Moments and cumulants of quadratic and bilinear forms of normally distributed variances. (In Russian) *Akademija Nauk Uzbekskoĭ SSR. Trudy Instituta Matematiki i Mehaniki im. V.I. Romanovskogo*, 15, 69-77.

Kawada, Y. (1950). Independence of quadratic forms in normally correlated variables. *The Annals of Mathematical Statistics*, 21, 614-615.

Kendall, M.G. (1971). Studies in the history of probability and statistics. XXVI. The work of Ernst Abbe. *Biometrika*, 58, 369-373.

Kendall, M.G. & Stuart, A. (1969). *The Advanced Theory of Statistics. Volume 1, Distribution Theory*. Third Edition. Charles Griffin & Company Ltd., London.

Khatri, C.G. (1959). On conditions for the forms of the type $\underset{\sim}{X}\underset{\sim}{A}\underset{\sim}{X}'$ to be distributed independently or to obey Wishart distribution. *Calcutta Statistical Association Bulletin*, 8, 162-168.

Khatri, C.G. (1961a). Cumulants and higher order uncorrelation of certain functions of normal variates. *Calcutta Statistical Association Bulletin*, 10, 93-98. [Errata: 12(1963), 67]

Khatri, C.G. (1961b). A simplified approach to the derivation of the theorems on the rank of a matrix. *Journal of the Maharaja Sayajirao University of Baroda*, 10, 1-5.

Khatri, C.G. (1962). Conditions for Wishartness and independence of second degree polynomials in a normal vector. *The Annals of Mathematical Statistics*, 33, 1002-1007.

Khatri, C.G. (1963). Further contributions to Wishartness and independence of second degree polynomials in normal vectors. *Journal of the Indian Statistical Association*, 1, 61-70.

Khatri, C.G. (1968). Some results for the singular normal multivariate regression models. *Sankhyā Series A*, 30, 267-280.

Koerts, J. & Abrahamse, A.P.J. (1969). *On the Theory and Application of the General Linear Model.* Rotterdam University Press, Rotterdam.

Laha, R.G. (1956). On the stochastic independence of two second-degree polynomial statistics in normally distributed variates. *The Annals of Mathematical Statistics,* 27, 790-796.

Laha, R.G. & Lukacs, E. (1960). On certain functions of normal variates which are uncorrelated of a higher order. *Biometrika,* 47, 175-176.

Lancaster, H.O. (1954). Traces and cumulants of quadratic forms in normal variables. *Journal of the Royal Statistical Society Series B,* 16, 247-254.

Lancaster, H.O. (1969). *The Chi-squared Distribution.* John Wiley & Sons, Inc., New York.

Loynes, R.M. (1966). On idempotent matrices. *The Annals of Mathematical Statistics,* 37, 295-296.

Lukacs, E. (1970). *Characteristic Functions.* Second Edition, revised and enlarged. Charles Griffin & Company, Ltd., London.

Lukacs, E. & Laha, R.G. (1964). *Applications of Characteristic Functions.* Hafner Publishing Company, New York. [Number 14 of Griffin's Statistical Monographs and Courses. Charles Griffin & Company, Ltd., London]

Luther, N.Y. (1965). Decomposition of symmetric matrices and distributions of quadratic forms. *The Annals of Mathematical Statistics,* 36, 683-690.

Madow, W.G. (1938). Contributions to the theory of multivariate statistical analysis. *Transactions of the American Mathematical Society,* 44, 454-495.

Madow, W.G. (1940). The distribution of quadratic forms in non-central normal random variables. *The Annals of Mathematical Statistics,* 11, 100-103.

Mäkeläinen, T. (1966). On quadratic forms in normal variables. *Societas Scientiarum Fennica. Commentationes Physico-Mathematicae,* 31 (12), 1-6.

Mann, H.B. (1949). *Analysis and Design of Experiments.* Dover Publications, Inc., New York.

Marcus, M. (1960). Basic theorems in matrix theory. U.S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series, 57. [Reprinted with Addenda and Errata, 1964]

Matérn, B. (1949). Independence of non-negative quadratic forms in normally correlated variables. *The Annals of Mathematical Statistics*, 20, 119-120.

Matusita, K. (1949). Note on the independence of certain statistics. *Annals of the Institute of Statistical Mathematics Tokyo*, 1, 79-82.

Mirsky, L. (1955). *An Introduction to Linear Algebra.* Oxford University Press, Oxford.

Mitra, S.K. (1968). A new class of g-inverse of square matrices. *Sankhyā Series A*, 30, 323-330.

Mitra, S.K. & Bhimasankaram, P. (1970). Some results on idempotent matrices and a matrix equation connected with the distribution of quadratic forms. *Sankhyā Series A*, 32, 353-356.

Mitra, S.K. & Rao, C.R. (1968). Simultaneous reduction of a pair of quadratic forms. *Sankhyā Series A*, 39, 312-322.

Nelder, J.A. (1951). A note on the statistical independence of quadratic forms in the analysis of variance. *Biometrika*, 38, 482-483.

Ogasawara, T. & Takahashi, M. (1951). Independence of quadratic quantities in a normal system. *Journal of Science of the Hiroshima University Series A*, 15, 1-9.

Ogawa, J. (1946a). On a proof of the fundamental lemma in H. Sakamoto's paper on the independence of statistics. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 1 (15), 25-28.

Ogawa, J. (1946b). On the independence of statistics of quadratic forms. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 2, 98-111. [Translation: 3 (1947), 137-151]

Ogawa, J. (1948). On the independence of statistics between linear forms, quadratic forms and bilinear forms from normal distributions. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 4, 1-40. [Translation: *Annals of the Institute of Statistical Mathematics Tokyo*, 1 (1949), 83-108]

Ogawa, J. (1949). Independence of statistics of quadratic forms for non-central normal distribution. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 5, 51-55. [Translation: *Osaka Mathematical Journal*, 2 (1950), 151-159]

Olkin, I. (1957). An exposition on the independence of quadratic, bilinear and linear forms for normal variates. Dittoed notes, 12pp. [Addenda: 3pp.]

Patnaik, P.B. (1949). The non-central χ^2 and F-distributions and their applications. *Biometrika*, 36, 202-232.

Pearson, K. (1960). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*, 50, 157-175. [Reprinted: *Early Statistical Papers (Pearson Reprints)*, 9 (1956), 443-475]

Plackett, R.L. (1960). *Regression Analysis.* Oxford University Press, Oxford.

Pringle, R.M. & Rayner, A.A. (1970). *Generalized Inverse Matrices with Applications to Statistics.* Charles Griffin & Company, Ltd., London.

Rao, C.R. (1962). A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *Journal of the Royal Statistical Society Series B*, 24, 152-158.

Rao, C.R. (1965). *Linear Statistical Inference and Its Applications.* John Wiley & Sons, Inc., New York.

Rao, C.R. (1966). Generalized inverse for matrices and its applications in mathematical statistics. *Research Papers in Statistics. Festschrift for J. Neyman.* Edited by F.N. David. John Wiley & Sons, Ltd., London, 263-279.

Rao, C.R. & Mitra, S.K. (1971). *Generalized Inverse of Matrices and Its Applications.* John Wiley & Sons, Inc., New York.

Rao, C.R. & Mitra, S.K. (1972). Generalized inverse of a matrix and its applications. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 601-620.

Rayner, A.A. (1966). On the distribution of Pearson's chi-squared criterion. *South African Journal of Agricultural Science*, 9, 1029-1031.

Rayner, A.A. & Livingstone, D. (1965). On the distribution of quadratic forms in singular normal variates. *South African Journal of Agricultural Science*, 8, 357-369.

Rayner, A.A. & Nevin B.M. (1970). Conditions for a quadratic form in normal variates to have chi-squared distribution. *Agrochemophysica*, 2, 93-94.

Rohde, C.A.; Urquhart, N.S., & Searle, S.R. (1966). Quadratic forms in singular normal variables. Paper No. BU-127, Biometrics Unit, Cornell University, 15pp.

Sakamoto, H. (1944a). Problems on the independence of statistics and degrees of freedom. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 1 (3), 65-74.

Sakamoto, H. (1944b). On the independence of two statistics. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo,* 1 (9), 1-25.

Sakamoto, H. (1946). On the independence of statistics and the judgement of degree of freedom in the theory of analysis of variance. (In Japanese) *Research Memoirs of the Institute of Statistical Mathematics Tokyo,* 2, 155-162. [Translation: *Annals of the Institute of Statistical Mathematics Tokyo,* 1 (1949), 109-122]

Sakamoto, H. (1948). On the independence and degrees of freedom of the maximum likelihood estimates. (In Japanese) *Bulletin of Mathematical Statistics,* 2 (1), 36-37; (2), 24-27.

Sakamoto, H. (1949). On independence of statistical quantities. (In Japanese) *Sûgaku,* 1, 263-274.

Searle, S.R. (1966). *Matrix Algebra for the Biological Sciences.* John Wiley & Sons, Inc., New York.

Searle, S.R. (1971). *Linear Models.* John Wiley & Sons, Inc., New York.

Searle, S.R.; Urquhart, N.S., & Evans, D.A. (1970). Cumulants of quadratic and bilinear forms in singular normal variables. Paper No. BU-205, Biometrics Unit, Cornell University, 15pp.

Seber, G.A.F. (1966). *The Linear Hypothesis: A General Theory.* Charles Griffin & Company, Ltd., London.

Shanbhag, D.N. (1966). On the independence of quadratic forms. *Journal of the Royal Statistical Society Series B,* 28, 582-583.

Shanbhag, D.N. (1968). Some remarks concerning Khatri's result on quadratic forms. *Biometrika,* 55, 593-596.

Shanbhag, D.N. (1970). On the distribution of a quadratic form. *Biometrika,* 57, 222-223.

Sjögren, P. (1967). Ett matrislemma till Cochrans sats. *Nordisk Matematisk Tidskrift,* 15, 157-158.

Smiley, M.F. (1964). The spectral theorem for finite matrices and Cochran's theorem. *The Annals of Mathematical Statistics,* 35, 443-444.

Styan, G.P.H. (1970). Notes on the distribution of quadratic forms in singular normal variables. *Biometrika,* 57, 567-572.

Styan, G.P.H. (1971). Personal communication.

Taussky, O. (1958). On the matrix theorem of A.T. Craig and H. Hotelling. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series A. Mathematical Sciences,* 61, 139-141.

Turin, G.L. (1960). The characteristic function of hermitian quadratic forms in complex normal variables. *Biometrika*, 47, 199-201.

Urquhart, N.S. & Searle, S.R. (1966). On the distribution of quadratic forms. Paper No. BU-212-M, Biometrics Unit, Cornell University, 11pp.

Wilks, S.S. (1962). *Mathematical Statistics*. John Wiley & Sons, Inc., New York.

Wishart, J. & Bartlett, M.S. (1932). The distribution of second order moment statistics in a normal system. *Proceedings of the Cambridge Philosophical Society*, 28, 455-459.

Zelen, M. (1962). Linear estimation and related topics. *Survey of Numerical Analysis*, McGraw-Hill Book Company, New York, 558-584.

Zelen, M. & Federer, W.T. (1965). Application of the calculus for factorial arguments. II. Analysis of factorials with unequal numbers of observations. *Sankhyā Series A*, 27, 383-400.