THE CRAIG-SAKAMOTO THEOREM

MYLÈNE FANNY DUMAIS

Department of Mathematics and Statistics, McGill University
805 rue Sherbrooke ouest, Montréal (Québec), Canada H3A 2K6

February 2, 2000

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

0-612-64348-4

THE CRAIG-SAKAMOTO THEOREM

MYLÈNE FANNY DUMAIS

Abstract

This thesis reviews the works that most influenced the progress and the development of the Craig-Sakamoto Theorem. This important theorem characterizes the independence of two quadratic forms in normal variables. We begin with a detailed and possibly complete outline of the history of this theorem, as well as several (correct) proofs published over the years. Furthermore, some misleading (or incorrect) proofs are reviewed and their lacunae explained. We conclude with a comprehensive bibliography on the Craig-Sakamoto Theorem; some associated references are also included.

Résumé

Cette thèse réunit les résultats qui ont le plus influencé l'avancement et le développement du théorème de Craig et Sakamoto. Ce théorème donne les conditions nécessaires et suffisantes pour obtenir l'indépendance de formes homogènes quadratiques. Dans le but de clarifier certaines croyances, nous débuterons avec une histoire détaillée et complète de ce théorème ainsi que plusieurs preuves publiées au fil des années. Outre cela, certaines preuves, qui ont été plus ou moins satisfaisantes, sont révisées et les erreurs (s'il y a lieu) qu'elles contiennent y sont expliquées. Finalement, nous concluons avec une bibliographie contenant plusieurs références sur le théorème de Craig et Sakamoto ainsi que d'autres sujets reliés à ce théorème.

Acknowledgements

I am extremely grateful to Professor George P. H. Styan; his suggestions and support were essential to the realization of this thesis. Moreover, I am indebted to Issie Scarowsky, whose 1973 McGill MSc thesis [203] covers various topics related to the distribution of quadratic forms in normal variables; indeed my thesis builds heavily on Chapters 2 and 4 and the bibliography of [203].

Special thanks go to Professors Xiao-Wen Chang (McGill University) and Junjiro Ogawa (Tokyo, Japan), whose detailed comments on the initial version of this thesis have greatly improved this final version. Professor Ogawa kindly made available to me an updated English translation of his 1993 paper [165] in Japanese and a copy of his 1997 joint paper [166] in English with Ingram Olkin. Professors Michael F. Driscoll, Chi-Kwong Li, Chang-yu Lu, Kameo Matusita, and Bao-Xue Zhang also gave me very helpful suggestions.

I am particularly grateful to Professor S. W. Drury (McGill University) for providing me with a proof of "Laha's Lemma" and for his permission to include this proof in my thesis. I am also indebted to Professor Chi-Kwong Li (College of William and Mary, Williamsburg, Virginia) for his permission to include the main result in his paper [117], and to Professors Li and Styan for their permission to include the main results in their joint paper [118], in this thesis.

I would also like to thank Ka Lok Chu, Kai-Tai Fang, William R. Gundberg, Jr., Shane T. Jensen, Serge Provost, Hong-Bin Qiu, Shayle R. Searle, Evelyn Matheson Styan, and Shagufta A. Sultan for their help, and for bringing several references to my attention. I also thank my family Gaston, Bibiane and Emilie Dumais for their continual moral support and Philippe Daigneault for his love and support.

Many thanks go to Robert V. Hogg (University of Iowa) for providing me with the portrait of Allen Thornton Craig (reproduced on page v, see also [227]), and to Jaap Hartman (International Statistical Institute) for providing me with the jpeg file of the portrait of Heihachi Sakamoto (reproduced on page v, see also [79]), and to Daniel Berze and Willem van Zwet for their help.

Some of the results in this thesis were presented by me at The Eighth International Workshop on Matrices and Statistics, Tampere, Finland, 6–7 August 1999, and by Professor Styan in the "Special Session on the Interaction Between Statistics and Matrix Theory" at the Annual Meeting of the Statistical Society of Canada, Regina, Saskatchewan, 7–8 June 1999, and at the Conference on Functional Analysis and Linear Algebra, Indian Statistical Institute–Delhi Centre, 3–7 January 2000. A related bibliography by Dumais and Styan [45] was prepared for The Seventh International Workshop on Matrix Methods for Statistics, in Celebration of T. W. Anderson's 80th Birthday (Fort Lauderdale, Florida, December 1998).

Table of Contents

Allen Thornton Craig
(1904–1978)

Heihachi Sakamoto
(b. 1914)

Chapter 1

Introduction and Preliminaries

Many papers have been written about the independence of two homogeneous quadratic forms in normal variables; some are misleading (or incorrect), some are correct but not easily understood, and some duplicate results published previously.

During the Second World War, Allen Thornton Craig (1904–1978)[1] the "Craig" of the famous book by Hogg and Craig [72], and Heihachi Sakamoto (b. 1914)[2] unaware of each other's work, both proposed a characterization for the stochastic independence of two homogeneous quadratic forms in independent normal variables with all means zero. This characterization was much simpler than the one proposed earlier in [30] by Cochran (1934). Craig's paper [35] was published in *The Annals of Mathematical Statistics* in 1943, while Sakamoto's paper [196] was presented at a "Lecture at the Annual Math-Physics Meeting [in Tokyo, Japan] on July 19, 1943" and published (in Japanese) in 1944. We will refer to this characterization as the Craig-Sakamoto Theorem. After the Second World War, the Craig-Sakamoto Theorem was extended in various ways and is now used to characterize the independence of two second-degree polynomials in normal variables not necessarily mutually independent and with means not necessarily zero.

In this chapter we review the underlying theorems and lemmas required to understand the Craig-Sakamoto Theorem. We cover basic information on matrices and on random vectors following a multivariate normal distribution, as well as the theorems needed in the different proofs of the Craig-Sakamoto Theorem. In Chapter 2, we present the views of various authors on the facts related to the Craig-Sakamoto Theorem, and in Chapter 3 we examine the facts and the proofs that have been published; in Chapter 4 we cover recent developments. Our thesis builds upon Chapters 2 and 4 of the 1973 McGill MSc thesis by Scarowsky [203], as well as the 1984 MA thesis [62] by Gundberg and the more recent work by Driscoll and Gundberg [42], Ogawa [165], and Ogawa and Olkin [166].

---

[1]Allen Thornton Craig was born in 5 August 1904 and died in 27 November 1978. For more biographical information and a photograph see [227].

[2]Heihachi Sakamoto was born on 16 August 1914 and elected to membership in the International Statistical Institute (ISI) in 1976; a portrait is available in the ISI Portrait Collection, cf. [79].

In preparation for our discussion of the Craig-Sakamoto Theorem, we first introduce the notation to be used and some preliminary results on matrices. We will denote matrices by capital letters and column vectors by lower case letters. All our matrices and vectors will be real (unless stated to the contrary).

The transpose A' of the $m \times n$ matrix A is the $n \times m$ matrix that has its rows equal to the corresponding columns of A; row vectors will always be primed. If B is an $m \times p$ matrix then we write $(A : B)$ for the $m \times (n+p)$ partitioned matrix with A placed next to B. The rank of a matrix A, denoted by $\text{rank}(A)$, is the dimension of its column space or its row space, while for the square matrix A we write $|A|$ or $\det(A)$ for its determinant, $\text{tr}(A)$ for its trace, and A^p for its pth power. A square matrix A is said to be symmetric when $A = A'$ and idempotent when $A = A^2$; an idempotent matrix need not be symmetric.

If there exists a matrix B such that $AB = BA = I_n$, the $n \times n$ identity matrix, then A is square and said to be nonsingular and B is its inverse denoted A^{-1}. If A is an $n \times n$ matrix, then A being nonsingular (or invertible) is equivalent to having $\text{rank}(A) = n$. The square matrix P is said to be orthogonal when $P' = P^{-1}$.

The set of eigenvalues $\{\text{ch}(A)\}$ of a square matrix A is the set of scalars $\{\lambda\}$ such that $Ax = \lambda x$; the associated nonnull vectors x are called eigenvectors and $|\lambda I_n - A|$ is called the characteristic polynomial, with its roots equal to the eigenvalues of A. If A is an $n \times n$ symmetric matrix, then all its eigenvalues are real, and we write $\text{ch}_i(A)$ for the ith largest eigenvalue. For any $n \times n$ symmetric matrix A with $\text{rank}(A) = r \leq n$, there exists an orthogonal matrix P such that $P'AP = D$, an $n \times n$ diagonal matrix containing the r nonzero eigenvalues on its diagonal and all other elements zero.

We also have the following fundamental equalities:

$$\begin{aligned} \text{tr}(A) &= \sum_{j=1}^{n} \text{ch}_j(A) \\ \det(A) = |A| &= \prod_{j=1}^{n} \text{ch}_j(A). \end{aligned}$$

Let A be an $m \times n$ matrix and let B be $n \times m$, so the products AB and BA are both defined. Then:

$$\begin{aligned} \{\text{nonzero ch}(AB)\} &= \{\text{nonzero ch}(BA)\} \\ \text{tr}(AB) &= \text{tr}(BA) \\ |I_m - AB| &= |I_n - BA|. \end{aligned}$$

A symmetric matrix A is positive definite (pd) when $x'Ax > 0$ for all $x \neq 0$, and nonnegative definite (nnd) when $x'Ax \geq 0$ for all x. A symmetric matrix A is positive definite if and only if all of its eigenvalues are strictly positive, and nonnegative definite if and only if all of its eigenvalues are nonnegative. If the $n \times n$ symmetric matrix A is nonnegative definite with rank r, then there exists an $n \times r$ matrix T of full column rank r, so that $A = TT'$. Since the trace $\text{tr}(A) = \text{tr}(TT')$ equals the sum of squares of the elements in T, we see that $\text{tr}(A) = \text{tr}(TT') = 0 \Leftrightarrow T = 0 \Leftrightarrow A = 0$.

We now introduce in detail our main topic of interest. Let $x = (x_1, ..., x_n)'$ be a random vector that follows a multivariate normal distribution. Then the quadratic expression $q = x'Ax$ is a "quadratic form in normal variables"; here $A = \{a_{ij}\}$ is an $n \times n$ nonrandom matrix. Another way to write q is:

$$q = x'Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j. \tag{1.1}$$

We (may and will) always choose the matrix A to be symmetric, since $q = x'Ax = (x'Ax)' = x'A'x = x'\{\frac{1}{2}(A + A')\}x$, and the matrix $\frac{1}{2}(A + A')$ is symmetric.

The quadratic form is called nonhomogeneous if its quadratic expansion contains quadratic, linear, and constant terms such as the polynomial $x'Ax + b'x + c$. A bilinear form is the sum of crossproducts between two distinct random vectors: $x'By$; here x is $m \times 1$, y is $n \times 1$, and B is $m \times n$.

The $n \times 1$ random vector x follows a multivariate normal distribution with mean vector $\mathrm{E}(x) = \mu = \{\mu_i\}$ and nonnegative definite dispersion matrix $\mathrm{D}(x) = V = \{v_{ij}\}$ whenever the scalar product $a'x$ follows a univariate normal distribution with mean $a'x$ and variance $a'Va$, for every $n \times 1$ vector a. We then write $x \sim \mathrm{N}(\mu, V)$ and have the following results:

$$\begin{aligned} x_j &\sim \mathrm{N}(\mu_j, v_{jj}); \quad j = 1, \ldots, n; \\ \mathrm{E}\{\exp(sx_j)\} &= \exp\{s\mu_j - \tfrac{1}{2}s^2 v_{jj}\}; \\ \mathrm{E}\{\exp(t'x)\} &= \exp\{t'\mu - \tfrac{1}{2}t'Vt\}. \end{aligned}$$

When the dispersion matrix V is positive definite (and thus nonsingular) then we have the following probability density functions:

$$\begin{aligned} \mathrm{pdf}(x) &= (2\pi)^{-n/2}|V|^{-1/2}\exp\left\{-\frac{1}{2}(x-\mu)'V^{-1}(x-\mu)\right\} \\ \mathrm{pdf}(x_j) &= (2\pi v_{jj})^{-1/2}\exp\left\{-\frac{(x_j-\mu_j)^2}{2v_{jj}}\right\}; \quad j = 1, \ldots, n. \end{aligned}$$

Our first theorem gives the characteristic function of a homogeneous quadratic form; we believe this version was first established in 1966 by Mäkeläinen [130]. To prove this theorem we note that if the moment generating function $M(s) = \mathrm{E}\{exp(sy)\}$ of a random variable y exists in the region $|s| \leq \varepsilon$, with $\varepsilon > 0$, then the characteristic function $\phi(s) = M(is)$ for all real s, cf. e.g., Lukacs [123], p. 11 and §7.1; here $i = \sqrt{-1}$.

Theorem 1 (The characteristic function of $x'Ax$) *If A is a real symmetric $n \times n$ nonrandom matrix and $x \sim \mathrm{N}(\mu, V)$, with V nonnegative definite, then $x'Ax$ has the characteristic function:*

$$\mathrm{E}\{\exp(isx'Ax)\} = \frac{\exp\{is\mu'(I - 2isAV)^{-1}A\mu\}}{|I - 2isAV|^{1/2}}. \tag{1.2}$$

To prove this theorem, we follow the presentation in Chapter 2 of Scarowsky [203] and use the following lemmas:

Lemma 2 *If $x \sim \mathrm{N}_n(\mu, V)$, then for any real $m \times n$ matrix H, the $m \times 1$ vector $Hx \sim \mathrm{N}_m(H\mu, HVH')$.*

The proof follows immediately from the definition of the multivariate normal distribution.

Lemma 3 *Let the random variable $z \sim \mathrm{N}(0,1)$. Then the joint moment generating function of z^2 and z is:*

$$\mathrm{E}\{\exp(sz^2 + tz)\} = \frac{\exp\left\{\frac{t^2}{2(1-2s)}\right\}}{(1-2s)^{1/2}}, \quad s < \tfrac{1}{2}. \tag{1.3}$$

Proof. We have

$$\begin{aligned}\mathrm{E}\{\exp(sz^2 + tz)\} &= \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \exp\left\{\frac{-z^2}{2} + sz^2 + tz\right\} dz \\ &= \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(z^2(1-2s) - 2tz\right)\right\} dz.\end{aligned}$$

Now let us put $w = z(1-2s)^{1/2} - t(1-2s)^{-1/2}$, with $s < 1/2$ (and so w is real). Then $w^2 = z^2(1-2s) + t^2(1-2s)^{-1} - 2zt$ so that

$$\begin{aligned}\mathrm{E}\{\exp(sz^2 + tz)\} &= \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \frac{\exp\left\{\frac{t^2}{2(1-2s)}\right\}}{(1-2s)^{1/2}} \exp(-w^2/2) dw \\ &= \frac{\exp\left\{\frac{t^2}{2(1-2s)}\right\}}{(1-2s)^{1/2}}; \quad s < \tfrac{1}{2},\end{aligned}$$

since the integral here is the integral of the probability density function from $-\infty$ to ∞ of the standard (univariate) normal distribution and so is equal to 1. Our proof is complete. □

Lemma 4 *If B and C are two $m \times n$ matrices such that $I_n - C'B$ is nonsingular, then $I_m - BC'$ is nonsingular, and*

$$(I_m - BC')^{-1} = I + B(I_n - C'B)^{-1}C'. \tag{1.4}$$

Proof. We have, with I denoting either I_m or I_n, that

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

and our proof is complete. □

Our next lemma can be found in the classic book by Anderson [4] (First Edition: pp. 25–26; Second Edition: pp. 31–33).

Lemma 5 *Let the $n \times 1$ random vector $x \sim \mathrm{N}(\mu, V)$, where V is $n \times n$. If $\mathrm{rank}(V) = r$, then there exists an $r \times 1$ random vector $y \sim \mathrm{N}(0, I_r)$ such that $x = Ty + \mu$ and $TT' = V$.*

Proof. Since V is symmetric we may write $V = V' = PDP'$, where P is orthogonal and D diagonal, and

$$V = PDP' = (P_1 : P_2)\begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} P_1' \\ P_2' \end{pmatrix},$$

where $P = (P_1 : P_2)$ and D_r is an $r \times r$ positive definite diagonal matrix. Here P_1 is $n \times r$ and P_2 is $n \times (n-r)$.

We introduce the nonsingular matrix

$$G = \begin{pmatrix} D_r^{-1/2} & 0 \\ 0 & I_{n-r} \end{pmatrix} P'.$$

Then

$$GVG' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

We now let $w = Gx$; then by Lemma 2, $w = Gx \sim \mathrm{N}(G\mu, GVG')$. Write $w' = (w_1' : w_2')$, where w_1 is $r \times 1$ and w_2 is $(n-r) \times 1$ and write $\nu' = (G\mu)' = (\nu_1' : \nu_2')$, with ν_1' and ν_2' having the same dimensions as w_1' and w_2', respectively. Then $w_1 \sim \mathrm{N}_r(\nu_1, I_r)$ and $w_2 = \nu_2$ with probability 1. Let

$$G^{-1} = (T : S),$$

where T is $n \times r$ and S is $n \times (n-r)$. Hence

$$\begin{aligned} x &= G^{-1}w \\ &= Tw_1 + Sw_2 \\ &= T(w_1 - \nu_1) + T\nu_1 + S\nu_2 \\ &= T(w_1 - \nu_1) + G^{-1}\nu \\ &= Ty + \mu, \end{aligned}$$

where $y = w_1 - \nu_1 \sim \mathrm{N}_r(0, I)$. It follows that

$$V = G^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} G^{-1'} = (T : S)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} T' \\ S' \end{pmatrix} = TT'$$

and the lemma is proved. □

Proof of Theorem 1. Since $x \sim \mathrm{N}(\mu, V)$, there exists (cf. Lemma 5) an $r \times 1$ random vector $y \sim \mathrm{N}(0, I_r)$ such that $x = Ty + \mu$ and $TT' = V$. Thus

$$x'Ax = (Ty + \mu)'A(Ty + \mu) = y'T'ATy + 2\mu'ATy + \mu'A\mu. \tag{1.5}$$

Since $T'AT$ is symmetric, it may be expressed as $Q\Delta Q'$, where Q is orthogonal and Δ is diagonal with the eigenvalues of $T'AT$ on the diagonal. (We note that the matrices $T'AT$, ATT' and AV

all have the same nonzero eigenvalues.) Substituting and using $QQ' = I$ in the second term of (1.5), we obtain

$$x'Ax = y'Q\Delta Q'y + 2\mu'ATQQ'y + \mu'A\mu. \tag{1.6}$$

From Lemma 2, we see that $Q'y \sim \mathrm{N}(0, I_r)$, since Q is orthogonal. We now let $Q'y = z$ and replace $2\mu'ATQ$ by ν' and $\mu'A\mu$ by α to obtain

$$x'Ax = z'\Delta z + \nu'z + \alpha = \sum_{j=1}^{r} \delta_j z_j^2 + \sum_{j=1}^{r} \nu_j z_j + \alpha. \tag{1.7}$$

We now use (1.7) to find the moment generating of $x'Ax$,

$$\begin{aligned}
\mathrm{E}[\exp(sx'Ax)] &= \mathrm{E}\left[\exp\left\{s\left(\sum_{j=1}^{r}\delta_j z_j^2 + \sum_{j=1}^{r}\nu_j z_j + \alpha\right)\right\}\right] \\
&= e^{s\alpha}\mathrm{E}\left[\exp\left\{s\left(\sum_{j=1}^{r}\delta_j z_j^2 + \nu_j z_j\right)\right\}\right] \\
&= e^{s\alpha}\mathrm{E}\left[\prod_{j=1}^{r}\exp\{s(\delta_j z_j^2 + \nu_j z_j)\}\right] \\
&= e^{s\alpha}\prod_{j=1}^{r}\mathrm{E}[\exp\{s(\delta_j z_j^2 + \nu_j z_j)\}]. && (1.8)
\end{aligned}$$

in view of the independence of the z_j's. Using Lemma 3, we now obtain:

$$\mathrm{E}[\exp\{sx'Ax\}] = \exp(s\alpha)\prod_{j=1}^{r}\left[\frac{\exp\left\{\frac{s^2\nu_j^2}{2(1-2s\delta_j)}\right\}}{(1-2s\delta_j)^{1/2}}\right] \tag{1.9}$$

provided $s\delta_j < 1/2$ for all $j = 1, \ldots, r$; if at least one $\delta_j > 0$ then this condition is equivalent to $s < 1/(2\delta_{\max})$, where $\delta_{\max} = \max(\delta_j)$.

Looking at the denominator of (1.9), we see that the product $\prod_{j=1}^{r}(1 - 2s\delta_j)^{1/2}$ involves the product of eigenvalues and we know that the product of eigenvalues of a matrix equals its determinant, and so:

$$\begin{aligned}
\prod_{j=1}^{r}(1-2s\delta_j)^{1/2} &= \prod_{j=1}^{r}\mathrm{ch}_j(I - 2sT'AT)^{1/2} \\
&= \left[\prod_{j=1}^{r}\mathrm{ch}_j(I - 2sATT')\right]^{1/2} \\
&= |I - 2sATT'|^{1/2} && (1.10) \\
&= |I - 2sAV|^{1/2}. && (1.11)
\end{aligned}$$

Now, all that is left is the numerator of the right-hand side of (1.9):

$$\begin{aligned}\prod_{j=1}^{r} \exp\left\{\frac{s^2\nu_j^2}{2(1-2s\delta_j)}\right\} &= \exp\left\{\sum_{j=1}^{r}\frac{s^2\nu_j^2}{2(1-2s\delta_j)}\right\} \\ &= \exp\left\{\frac{s^2}{2}\nu'(I-2s\Delta)^{-1}\nu\right\}. \end{aligned} \tag{1.12}$$

Since $T'AT = Q\Delta Q'$, we have:

$$Q(1-2s\Delta)^{-1}Q' = [Q(I-2s\Delta)Q']^{-1} = [QQ'-2sQ\Delta Q']^{-1} = [I-2sT'AT]^{-1}. \tag{1.13}$$

Using $\nu' = 2\mu' ATQ$ and $Q(1-2s\Delta)^{-1}Q' = [I-2sT'AT]^{-1}$, we obtain:

$$\frac{s^2}{2}\nu'(I-2s\Delta)^{-1}\nu = 2s^2\mu' AT[I-2sT'AT]^{-1}T'A'\mu. \tag{1.14}$$

And so, with the help of Lemma 4, it follows that:

$$\begin{aligned}\exp(s\alpha)\prod_{j=1}^{r}\exp\left\{\frac{s^2\nu_j^2}{2(1-2s\delta_j)}\right\} &= \exp\left\{2s^2\mu' AT[I-2sT'AT]^{-1}T'A'\mu + s\mu' A\mu\right\} \\ &= \exp\left\{s\mu'(2sAT[I-2sT'AT]^{-1}T' + I)A\mu\right\} \\ &= \exp\left\{s\mu'(I-2sATT')^{-1}A\mu\right\} \\ &= \exp\left\{s\mu'(I-2sAV)^{-1}A\mu\right\}.\end{aligned}$$

Hence by replacing the numerator and the denominator and recalling that the characteristic function $f(s) = m(is)$, where $m(s)$ is the moment-generating function (mgf), Theorem 1 is established.
□

Lemma 6 *If the eigenvalues of a square matrix are all less than* 1 *in absolute value then*

$$(I-G)^{-1} = \sum_{h=0}^{\infty} G^h. \tag{1.15}$$

Proof. See. e.g., Mirsky [142, p. 332]. □

Lemma 7 *If the eigenvalues γ_h of a square matrix are all real and less than* 1 *in absolute value then*

$$\log(|I-G|) = -\sum_{j=1}^{\infty}\frac{1}{j}\mathrm{tr}(G^j). \tag{1.16}$$

Proof. We have

$$\begin{aligned}\log(|I-G|) &= \log\left[\prod_{h=1}^{n}(1-\gamma_h)\right] \\ &= \sum_{h=1}^{n}\log(1-\gamma_h)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{h=1}^{n}\left[-\sum_{j=1}^{\infty}\frac{\gamma_h^j}{j}\right] \\
&= -\sum_{j=1}^{\infty}\frac{1}{j}\left[\sum_{h=1}^{n}\gamma_h^j\right] \\
&= -\sum_{j=1}^{\infty}\frac{1}{j}\mathrm{tr}(G^j), \qquad (1.17)
\end{aligned}
$$

and the lemma is established. □

We end this chapter with the cumulant generating function and the cumulants of the quadratic form $x'Ax$. We believe that Dieulefait (1951) and Lancaster (1954) were the first to obtain independently this result for $x \sim \mathrm{N}(0, I)$. We present the cumulant generating function as stated by Khatri (1963), and then by Rohde, Urquhart, and Searle (1966), for $x \sim \mathrm{N}(\mu, V)$, with V possibly singular.

Theorem 8 (The cumulant generating function of $x'Ax$) *If x follows a multivariate normal distribution with mean μ and dispersion matrix V, with V possibly singular and if A is a symmetric matrix; then the cumulant generating function of $x'Ax$ is*

$$
\phi(s) = \sum_{j=1}^{\infty}\frac{s^j}{j!}\left\{j!2^{j-1}\left[\mu'(AV)^{j-1}A\mu + \frac{1}{j}\mathrm{tr}(AV)^j\right]\right\}. \qquad (1.18)
$$

Proof. We follow Scarowsky [203, Theorem 2.2]. Using Theorem 1, we express $\phi(s)$ as

$$
\phi(s) = \log[\mathrm{E}(\exp\{sx'Ax\})] = s\mu'(I-2sAV)^{-1}A\mu - \tfrac{1}{2}\log|I-2sAV|. \qquad (1.19)
$$

It is possible to find a positive number ε, such that for all $|s| < \varepsilon$ we have $|\mathrm{ch}(sAV)| < 1$. Using Lemma 6 and Lemma 7 gives

$$
\begin{aligned}
\phi(s) &= s\mu'\left\{\sum_{h=0}^{\infty}2^h s^h (AV)^h\right\}\mu + \frac{1}{2}\sum_{j=1}^{\infty}2^j s^j\left\{\frac{1}{j}\mathrm{tr}(AV)^j\right\} \\
&= \sum_{j=1}^{\infty}2^{j-1}s^j\mu'(AV)^{j-1}A\mu + \sum_{j=1}^{\infty}2^{j-1}s^j\left\{\frac{1}{j}\mathrm{tr}(AV)^j\right\} \qquad (1.20)
\end{aligned}
$$

and our proof is complete. □

It follows at once from Theorem 8 that the jth cumulant of $x'Ax$ is

$$
\mu'(AV)^{j-1}A\mu + \frac{1}{j}\mathrm{tr}(AV)^j; \quad j = 1, 2, \ldots \qquad (1.21)
$$

Chapter 2

Historical Points of View

The history of the Craig-Sakamoto Theorem has often been a source of disagreement among statisticians. Since so many papers were published in the 1940s, 1950s, and the 1960s in various places and in various languages, it was difficult for researchers to consult all the work that had been done on the topic before publishing their own results. Therefore, there exist many instances of duplication, variation, misconception, and lacunae, such as inevitably lead scholars to attempt to clarify the facts. To illustrate this, we present different versions of the development of the Craig-Sakamoto Theorem in this chapter.

In 1934, W. G. Cochran [30] introduced some corollaries on the chi-squareness and the independence of quadratic forms and proved the following result.

Theorem 9 (Cochran 1934 [30]) *If a random vector x is distributed as $\mathrm{N}(0, I)$ and if $A = A', B = B'$ are nonrandom, then the quadratic forms $x'Ax$ and $x'Bx$ are stochastically independent if and only if*

$$|I - sA - tB| = |I - sA| \cdot |I - tB| \qquad \forall \text{ real } s \text{ and } t. \tag{2.1}$$

Based on this result, Craig [34] noticed in 1938 that if $x'Ax$ and $x'Bx$ are independent, then

$$\mathrm{rank}(A + B) = \mathrm{rank}(A) + \mathrm{rank}(B), \tag{2.2}$$

with rank identifying the number of associated independent variables. Cochran [30] proved his finding by showing that if the joint moment generating function of $x'Ax$ and $x'Bx$ factorizes as the product of the moment generating functions of $x'Ax$ and of $x'Bx$, then independence was obtained and vice versa. Nevertheless, these two conditions were difficult to apply. Consequently, the challenge was to discover a nice and simple condition for independence that would make independence easy to verify. In 1943 Craig [35] and in 1944 Sakamoto [196] asserted that (2.1) holds if and only if $AB = 0$; their proofs, however, were incomplete.

After the discovery of this result, many researchers then tried to produce a complete proof for the simple central case with dispersion matrix I and then for the general case $\mathrm{N}(\mu, V)$, first with V

positive definite and then with V possibly singular. On several occasions, various scholars tried to distinguish the correct proofs from the misleading ones. The first survey was Scarowsky's 1973 MSc thesis entitled *Quadratic Forms in Normal Variables* [203]. His work includes the characteristic function of a quadratic form, the conditions for chi-squaredness and independence. With respect to the Craig-Sakamoto Theorem, Scarowsky noted that

- Craig (1943), Sakamoto (1944), Hotelling (1944), and Ogawa (1946) were unable to provide a complete proof of the Craig-Sakamoto Theorem.
- The first to provide a complete proof was Matusita (1949) followed by Aitken (1950): both proofs covered the case where the dispersion matrix V is positive definite.
- Carpenter (1950) and Ogawa (1950) were the first to extend the result to the non-central case.

Moreover, Scarowsky's thesis contains Ogasawara and Takahashi's (1951) proofs for the simplest case with $x \sim \mathrm{N}(0, I)$ and the most general case with $x \sim \mathrm{N}(\mu, V)$, with V possibly singular. Scarowsky also included proofs for the independence of bilinear forms and the mutual independence of more than two quadratic forms, and he listed more than a hundred references on the subject.

In 1984, William R. Gundberg Jr. completed a Master's thesis [62] entitled *A History of Results on Independence of Quadratic Forms in Normal Variates* and in 1986 published a paper [42] on his findings with his thesis adviser Michael F. Driscoll. Apparently not satisfied with the proofs available to them, they wanted to provide a more complete proof and a further description of the history of the Craig-Sakamoto Theorem. Driscoll and Gundberg [42] also stated several problems one may encounter in various treatments of the theorem.

As Driscoll and Gundberg [42, p. 65] noted,

> "The history of Craig's theorem is not a happy one. The authors of the earlier articles in the literature tended to make errors of a linear-algebraic nature. Authors of more recently published textbooks have given incorrect or misleadingly incomplete coverage of Craig's theorem and its proof."

Driscoll and Gundberg [42] reviewed all possible cases from the simplest case with $x \sim \mathrm{N}(0, I)$ to the most general case with $x \sim \mathrm{N}(\mu, V)$, with V possibly singular, and observed the following:

- for the $\mathrm{N}(0, I)$ case:
  - Cochran (1934) [$|I - 2sA| \cdot |I - 2tB| = |I - 2sA - 2tB|$]
  - Craig (1938) [first to observe rank additivity: $\mathrm{rank}(A + B) = \mathrm{rank}(A) + \mathrm{rank}(B)$]
  - Craig (1943) [first to state $AB = 0$, but proof incomplete]
  - Hotelling (1944) [incomplete proof (subtle gap)]
  - Craig (1947) [proved that two bilinear forms are independent iff $AB = 0$]
  - Lancaster (1954) [correct linear-algebraic proof]

- for the $N(0, V)$ case:
  - Ogawa (1949) [first correct linear-algebraic proof]
  - Aitken (1950) [first to transform $x \to V^{-1/2}x$]
  - Ogasawara and Takahashi (1951) [first (complete) treatment with V possibly singular]
  - Lancaster (1954)
- for the $N(\mu, I)$ and the $N(\mu, V)$ case
  - Carpenter (1950) [used $x \to V^{-1/2}x$, but his proof of $N(\mu, I)$ is incomplete]
  - Ogasawara and Takahashi (1951) [first (complete) treatment with V possibly singular]
  - Laha (1956) ["Laha's Lemma" stated but not proved]
  - Searle (1984) [partially completes Laha's proof].

All in all, Driscoll and Gundberg (1986) claimed they found no source that contained a correct, complete and detailed proof of the general case, and they observed that the information in the existing textbooks on the proof or on the history of the Craig-Sakamoto Theorem is often inadequate. Finally, they offered recommendations to future authors on how to treat the subject.

Two years later, Reid and Driscoll (1988) [192] revised the 1986 paper of Driscoll and Gundberg [42], following some research done by Reid (1986). According to Reid and Driscoll (1988)

- the independence of $x'Ax$ and $x'Bx$ does not imply $AVB = 0$ for all μ, only for $\mu = 0$. Krafft (1978) shows that it implies that AVB is only skew-symmetric.
- Krafft's proof can be considerably shortened using Zielinski's (1985) argument.
- Laha's proof and Driscoll's supplement both appear in Ogawa (1950).
- the claim of necessity can be proved using $\text{cov}(x'Ax, x'Bx) = 0$ which implies independence when $\mu = 0$ and A and B are nonnegative definite, as proposed by Matérn (1949) and Kawada (1950).

Furthermore, Reid and Driscoll (1988) included an elegant proof of the general case using cumulants which will be discussed in Chapter 3.

Even if these two last papers give a good historical review on the past events, the absence of references to the papers by Sakamoto [197, 200], Matusita [138], Ogasawara and Takahashi [157], and other Japanese statisticians who have produced key papers on this theorem is regrettable. To resolve this matter, Ogawa (1993) [165] presented

> " a fair description of the development of Craig-Sakamoto's theorem ... giving due credits to Japanese authors who have been overlooked by Western authors of papers and textbooks."

Ogawa had written three papers in 1946 [159], 1949 [162], and 1950 [164] on this topic; and in 1993, he expressed much disappointment in general about other authors' misunderstanding. In an effort to achieve the same aims as Driscoll and Gundberg (1986), Ogawa (1993) gave a complete listing of the Japanese accomplishments which he compared with those of other statisticians. The following is Ogawa's summary:

- The $N(0, I)$ case
  - was conjectured by Craig (1943) and Sakamoto (1944).
  - was proved by Matusita (1949), Ogasawara and Takahashi (1951), Lancaster (1954), and Mathai and Provost (1992).
  - was not completely proved by Craig (1943), Hotelling (1944), and Aitken (1950).
- The $N(\mu, V)$ case
  - was proved by Ogawa (1950), Laha (1956) [incomplete], Driscoll and Gundberg (1986), and Reid and Driscoll (1988).
  - was not completely proved by Carpenter (1950), Ogasawara and Takahashi (1951), Kendall and Stuart (1969), Johnson and Kotz (1970), Searle (1971), Mathai and Provost (1992).
- The proofs given by Ogawa (1949) and Kawada (1950) were different from the ones listed above, but they are complete. However, the ones proposed by Ogawa (1946) and Zielinski (1985) are incomplete in reasoning.

Ogawa (1993) also presented four different proofs, all by Japanese statisticians, for the central case,

- Matusita (1949)[1]
- Ogasawara and Takahashi (1951)
- Kawada (1950)
- an improvement of Nabeya's 1949 proof[2].

Moreover, Ogawa [165] criticized "Western statisticians" for overlooking the Japanese contributions to the development of the Craig-Sakamoto Theorem. Ogawa confirmed that Matusita (1949) was the first to give a full and correct proof, and not Lancaster (1954), as implied by Hogg and Craig [72]. In the description of the theorem by Johnson and Kotz (1970), Ogawa noticed that there was no mention at all of any Japanese involvement. Then, he reviewed Hotelling's 1944 proof and maintained that it is incomplete. Finally, he pointed out some lacunae in Mathai and Provost's 1992 book *Quadratic Forms in Random Variables* [135].

In 1997, in a joint paper [166] in English, Ogawa and Olkin critically examined the literature and revised some of Ogawa's 1993 opinions. Before presenting proofs for $N(0, I)$, they now declared that (contrary to Ogawa's belief in 1993) Aitken (1950) gave the third correct proof after Matusita and Ogawa, and that it is this proof which was given by Lancaster (1954). Concerning the description of the historical and mathematical facts, Ogawa and Olkin [166] maintained that the historical accounts provided in the books by Hogg and Craig [72], Searle [209], Guttman [65], and Hocking [70] were limited. Ogawa and Olkin considered that while the 1992 book by Mathai and Provost [135] may give many more details on the topic, they still found the proof of the non-central case and

[1] Ogawa (1993) claimed that Lancaster's 1954 proof is the same as Matusita's.

[2] Published in the paper by Ogawa [162].

the history to be unsatisfactory. Ogawa referred to Scarowsky (1973) and Driscoll and Gundberg (1986) as accurate sources for historical details, even if he considered them to be incomplete in other respects.

Then, Ogawa and Olkin [166] treated the central case with the dispersion matrix V being different from I and gave historical details. To the ones already listed, they added the works of Good (1963, 1966), Shanbhag (1966), and Nagase and Banerjee (1976). Moreover, they explained three proofs for the general case: Ogawa (1950) and Laha (1956), Reid and Driscoll (1988) and Driscoll and Krasnicka (1995), and Olkin (1997). Furthermore, they discussed the general case for the second degree polynomial, bilinear forms and for multivariate versions.

Chapter 3

The Craig-Sakamoto Theorem: 1943–1996

Now, with all the documents we have at hand, we introduce the proofs and present the historical facts surrounding the Craig-Sakamoto Theorem. We attempt to discuss the proofs as provided chronologically, explaining them and pointing out their lacunae (if any) and to clarify the whole situation.

3.1. From 1943 to 1949

3.1.1. Craig (1943), Hotelling (1944)

In 1943, after studying quadratic and bilinear forms, Craig [35] was the first to state a condition for their independence other than the one found in 1934 by Cochran [30], cf. (3.1).

Theorem 10 (Craig 1943) *Let x be a random vector that follows a multivariate normal distribution with mean zero and dispersion matrix I. If $x'Ax$ and $x'Bx$ are two quadratic forms with A and B symmetric, then these quadratic forms are independent if and only if $AB = 0$.*

Proof of sufficiency.

Assuming from the right-hand side of equation (2.1) that $AB = 0$, we obtain

$$|I - sA - tB| = |I - sA - tB + stAB| = |I - sA| \cdot |I - tB|. \tag{3.1}$$

Hence, by Theorem 9, $x'Ax$ and $x'Bx$ are independent.

The difficulty in proving Theorem 10 is to show that if there is independence then $AB = 0$. Cochran had already observed in 1934 that two distributions were independent if and only if their joint moment generating function (mgf) is equal to the product of the two marginal mgf's, and this is a key step in many proofs that followed in the literature. Hence, assuming independence and that s could be put equal to t without loss of generality, Craig (1943) expressed the mgf's as products involving the characteristic roots of A, B, and $A + B$:

$$
\begin{aligned}
|I - sA| &= (1 - s\alpha_1)\cdots(1 - s\alpha_a) && (3.2)\\
|I - sB| &= (1 - s\beta_1)\cdots(1 - s\beta_b) && (3.3)\\
|I - s(A + B)| &= (1 - s\gamma_1)\cdots(1 - s\gamma_c), && (3.4)
\end{aligned}
$$

where $\alpha_i, \beta_j, \gamma_k$ are the nonzero eigenvalues of A, B, and $A + B$, respectively. Thus $a + b = c$ and since the matrices are all real and symmetric we must have rank additivity:

$$c = \text{rank}(A + B) = \text{rank}(A) + \text{rank}(B) = a + b. \tag{3.5}$$

By Theorem 9, Craig noted that the product of (3.2) and (3.3) was equal to (3.4) and hence, each term γ_k could be paired with either the term α_i or the term β_j for $k = 1, \ldots, c; i = 1, \ldots, a; j = 1, \ldots, b$. Next, he defined an orthogonal matrix L to diagonalize A and B simultaneously[1]. Craig claimed that

$$L'ALL'BL = \begin{pmatrix} D_A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & D_B \end{pmatrix} = 0 \tag{3.6}$$

and so $AB = 0$; that D_A and D_B have no overlapping nonzero eigenvalues follows at once from the rank additivity (3.5). Hence it follows that the set of all eigenvalues of $A + B$ is the union of the set of eigenvalues of A and of B.

Hotelling (1944) [73], commenting on Craig's proof, observed that:

> "The proof given that the condition is sufficient is adequate, but Craig's treatment of its necessity consists essentially in its assertion".

Observing that Theorem 10 could have a "wide usefullness", Hotelling included a proof in his paper. After sufficiency was shown, he assumed the independence of $x'Ax$ and $x'Bx$. As Craig [35] did, Hotelling [73] used an orthogonal matrix P to diagonalize A and he was careful to mention that P did not necessarily diagonalize B. He set

$$P'AP = D = \begin{pmatrix} D_A & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.7}$$

where D_A is the diagonal matrix containing the nonzero eigenvalues of A, and he set $P'BP = M$. Then he showed that

$$|I - sA| = |P(I - sD)P'| = |P| \cdot |I - sD| \cdot |P'| = |PP'| \cdot |I - sD| = |I - sD|, \tag{3.8}$$

[1]Such an orthogonal matrix exists only if A and B commute.

and similarly

$$|I - tB| = |I - tM| \tag{3.9}$$

$$|I - sA - tB| = |I - sD - tM| \tag{3.10}$$

for B and $A + B$, respectively. Hence, by Theorem 9 he obtained

$$|I - sD| \cdot |I - tM| = |I - sD - tM|.$$

Next, he split the matrix $M = M_1 + M_2$ where

$$M_1 = \begin{pmatrix} E & C \\ C' & 0 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & F \end{pmatrix}, \tag{3.11}$$

and where E has the same rank as A. Letting $z' = x'P'$, he set $Q_1 = x'Ax = z'PAP'z$, $Q_2 = x'Bx = z'PBP'z = z'(M_1 + M_2)z = z'M_1z + z'M_2z = Q_3 + Q_3'$. Because it was obvious that $DM_2 = 0$, all that was left to show was that $DM_1 = 0$. He claimed that in this case, the three quadratic forms Q_1, Q_3, Q_3' are independent by this argument. Since Q_1 and Q_3' do not have any variates in common and are independent, then Q_1 and Q_3' are independent. Moreover, if Q_1 and Q_2 are independent by assumption, it then follows that Q_1 is independent of $Q_2 - Q_3' = Q_3$. Therefore, $DM = 0$ and $PDMP' = PDPP'MP' = AB = 0$.

Unfortunately, this proof did not satisfy many; some would claim that the argument was false. As Driscoll and Gundberg (1986) explained,

> "... Hotelling's proof contains not a falsity, but a subtle gap. ... By relying on a later correct proof of [the Craig-Sakamoto Theorem], it can be shown that there are no counterexamples to Hotelling's statement ... , but this leaves open the question of how this independence might be proved to salvage Hotelling's proof [2]."

It was shown by Baksalary and Hauke (1984) that a requirement for Q_1 and Q_2 to be independent is that Q_1 must be independent of the pair $(Q_2 + Q_3, Q_3)$. However, Ogawa (1949, 1993) and Ogawa and Olkin (1997) claimed that this argument does not hold. So $DM_1 = 0$ could not be shown. To support this, Ogawa referred to a counterexample by Bernstein[3].

Meanwhile, Sakamoto (1944), made the same observation independently of Craig (1943) and published his result that covered not only the $\mathrm{N}(0, I)$ case but also the $\mathrm{N}(0, V)$ case where V is positive definite.

3.1.2. Sakamoto (1944), Ogawa (1946), Craig (1947)

Theorem 11 (Sakamoto 1944) *Let $x \sim \mathrm{N}(0, V)$, where V is a positive definite matrix, and let A and B be symmetric matrices. The quadratic forms $x'Ax$ and $x'Bx$ are independent if and only if $AVB = 0$.*

[2] Cf. Driscoll and Gundberg [42], p. 66.

[3] Cf. Kolmogoroff (1933), Section 5 "Unabhängigkeit", pp. 8–11; footnote, p. 10.

Sakamoto was the first to propose that the $N(0, V)$ case could be reduced to the $N(0, I)$ case using a transformation. However, in his attempt to show the $N(0, I)$ case, Sakamoto made the same mistake as Craig (1943) did by diagonalizing A and B simultaneously. Although, the Craig-Sakamoto Theorem remained technically unproven at this point, it is interesting to notice that some statisticians, such as Kac (1945), were using this result.

In 1946, a totally different proof was proposed by Ogawa [159]. He considered A and B to be linear transformations in a linear vector space $\mathcal{L}$ where the nullspace is the set $\mathcal{N}_A = \{x|Ax = 0\}$. Then, he diagonalized A and noted that the dimension of the set $\mathcal{M} = \{x|Ax \neq 0\}$ is equal to the number of nonzero eigenvalues of A and $\dim(\mathcal{N}_A)$ is the number of the eigenvalues equal to zero and similarly for B. Then using Theorem 9, he set $s = t = 1/\lambda$ to obtain

$$|\lambda I_n - A| \cdot |\lambda I_n - B| = \lambda^n |\lambda I_n - A - B|. \tag{3.12}$$

Ogawa (1946) then claimed that

$$\dim(\mathcal{N}_A) + \dim(\mathcal{N}_B) = n + p,$$

where p is the multiplicity of the zero eigenvalue of $A+B$. Unfortunately, such a result requires that the set of nonzero eigenvalues of $A + B$ be equal to the union of the sets of nonzero eigenvalues of A and B where these sets are disjoint. These conditions on the sets of eigenvalues are the key step to prove the Craig-Sakamoto Theorem. However, Ogawa (1946) continued his argument without proving his claim. Furthermore, S. Nabeya pointed out some errors in Ogawa's proof, about which he communicated with Ogawa[4].

While Craig (1947) published a paper on bilinear forms, making a mistake in proving their independence, Ogawa and Sakamoto were working on this topic with three other Japanese reseachers: Nabeya (mentioned earlier), Matusita and Sugawara. In 1949, three acticles by Matusita [138], Ogawa [162], and Sakamoto [200] were all published in the *Annals of the Institute of Statistical Mathematics Tokyo*, "received" respectively on June 2, June 10, and June 30, 1948[5].

Matusita (1949) gave the first correct proof of the Craig-Sakamoto Theorem for the $N(0, V)$ case with $V = I$ as a special case. In a footnote[6], Matusita (1949) wrote:

> "I had this result in 1943, independently of A. T. Craig when H. Sakamoto asked me about the independence of quadratic forms, and informed it to him. He and some other colleagues of mine have searched for other proofs of this theorem (esp. M. Sugawara and S. Nabeya) or applied it to various problems. On seeing these recent investigations and those of some others, it seems to me of some use to publish my proof at this stage."

[4] We are unaware of any publication by S. Nabeya about this matter or on any topic related to the Craig-Sakamoto Theorem.

[5] The papers by Ogawa and by Sakamoto were published originally in Japanese, respectively in 1948 and 1946, in the *Research Memoirs of the Institute of Statistical Mathematics Tokyo*. A history of the Institute of Statistical Mathematics Tokyo by Kameo Matusita appears in [139].

[6] On page 82 of [138]

3.1.3. Matusita's 1949 Proof for $\mathrm{N}(0, V)$, with V Positive Definite

Let $x = (x_1, ..., x_n)'$ be a vector of random variables from an n-variate normal population with dispersion matrix V positive definite. There exists an orthogonal matrix P such that $P'VP = D^2$, where D is a real diagonal matrix with nonnegative components and let $V = TT'$. Then, it follows at once from Theorem 1 that the quadratic form $x'Ax$ has its moment generating function equal to

$$f(z) = |I - 2sA^*|^{-1/2},$$

where I is the $n \times n$ identity matrix and $A^* = P'T'ATP$ is symmetric. So the independence of $x'Ax$ and $x'Bx$ is equivalent to the equation:

$$|I - (sA^* + tB^*)| = |I - sA^*| \cdot |I - tB^*| \qquad \forall \textit{ real } s \textit{ and } t. \tag{3.13}$$

where s and t are independent variables, and A^* and B^* represent the matrices $P'T'ATP$ and $P'T'BTP$, respectively. Obviously the equation (3.13) holds when $AVB = 0$ and thus it is a sufficient condition.

Assuming that the two quadratic forms $x'Ax$ and $x'Bx$ are independent, then equation (3.13) holds. Now let $\alpha_1, ..., \alpha_r$ and $\beta_1, ..., \beta_{r'}$ be the nonzero eigenvalues of A^* and B^*, respectively. There exists an orthogonal matrix U such that $U'A^*U = \mathrm{diag}(\alpha_1, ..., \alpha_r, 0, ..., 0) = D_{A^*}$, say. Writing the left-hand side of equation (3.13) gives the following:

$$\begin{vmatrix} 1 - s\alpha_1 - tc_{11} & \cdots & -tc_{1r} & -tc_{1,r+1} & \cdots & -tc_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -tc_{r1} & \cdots & 1 - s\alpha_r - tc_{rr} & -tcr, r+1 & & -tc_{rn} \\ -tc_{r+1,1} & \cdots & -tc_{r+1,r} & 1 - tc_{r+1,r+1} & \cdots & tc_{r+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -tc_{n1} & \cdots & -tc_{nr} & -tc_{n,r+1} & \cdots & 1 - tc_{nn} \end{vmatrix}. \tag{3.14}$$

where $\{c_{ij}\}$ denotes the matrix $U'B^*U = C$. The coefficient of s^r is

$$\alpha_1 \cdots \alpha_r \begin{vmatrix} 1 - tc_{r+1,r+1} & \cdots & -tc_{r+1,n} \\ \vdots & \ddots & \vdots \\ 1 - c_{n,r+1} & \cdots & 1 - tc_{nn} \end{vmatrix} = (-1)^r \alpha_1 \cdots \alpha_r |I - tC_{22}|, \tag{3.15}$$

say, and it must be equal to $(-1)^r \alpha_1 \cdots \alpha_r |I - tB^*|$ for equation (3.13) to hold. From this we can infer that the nonzero eigenvalues of

$$C_{22} = \begin{pmatrix} c_{r+1,r+1} & \cdots & c_{r+1,n} \\ \vdots & \ddots & \vdots \\ c_{n,r+1} & \cdots & c_{nn} \end{pmatrix} \tag{3.16}$$

are $\beta_1, ..., \beta_{r'}$, that is, the same nonzero eigenvalues as B^*. Diagonalizing C_{22} using an $(n-r) \times (n-r)$ orthogonal matrix V_1, we then set

$$V_2 = \begin{pmatrix} I_r & 0 \\ 0 & V_1 \end{pmatrix}. \tag{3.17}$$

Putting $W = UV_2$ gives

$$W'A^*W = \text{diag}(\alpha_1, ..., \alpha_r) \tag{3.18}$$

and

$$W'B^*W = \begin{pmatrix} b^*_{11} & \cdots & & \cdots & b^*_{n1} & 0 \\ \cdots & \ddots & & & \vdots & \vdots \\ \cdots & & \beta_1 & 0 & \cdots & \vdots \\ \cdots & & & \ddots & & \vdots \\ \cdots & & & & \beta_{r'} & \\ b^*_{n1} & & 0 & & \cdots & 0 \end{pmatrix}. \tag{3.19}$$

Now looking at the norm of B^* which we may choose as the square root of the trace of $B^{*'}B^*$, we notice that

$$\sum{}' |b^*_{ij}|^2 + \sum_{i=1}^{r'} |\beta_i|^2 = \sum_{i=1}^{r'} |\beta_i|^2, \tag{3.20}$$

where $\sum'$ represents the summation running on (ij) where either $i \leq r$ or $j \leq r$, since the norm is invariant under the unitary transformation. We must, therefore, have all b^*_{ij} with $i \leq r$ or $j \leq r$ equal to zero. Thus,

$$W'A^*WW'B^*W = WA^*B^*W = 0. \tag{3.21}$$

This implies that $A^*B^* = 0$ and hence

$$A^*B^* = P'T'ATPP'T'BTP = 0. \tag{3.22}$$

Therefore, we obtain $ATT'B = AVB = 0$. □

3.1.4. Ogawa's 1949 Proof for $\mathrm{N}(0, I)$

Ogawa (1949) explained in depth the algebraic conditions needed for his proof and introduced two key lemmas due to Nabeya, whom he thanked in his paper for his help[7].

[7] In a footnote (on page 107 of [162]) Ogawa writes: "I am indebted to Mr. M. Sugawara and Mr. S. Nabeya of the Institute of Statistical Mathematics for advices and criticism while this paper was being prepared." We are not aware of any papers by either M. Sugawara or S. Nabeya on the Craig-Sakamoto Theorem.

Lemma 12 (Nabeya's First Lemma) *Let A and B be two $n \times n$ real symmetric matrices and let their ranks be a and b, respectively, and furthermore, let the rank of $C = A + B$ be r_C. If the rank-additivity relation $a + b = c$ holds, then the subspace of C generated by its eigenvectors is the direct sum of the subspaces of A and B which is generated by the eigenvectors of A and B. Moreover, if C is idempotent, then $AB = BA = 0$ and A and B are idempotent themselves.*

The last part of Lemma 12 was already proved by Cochran (1934) and is the special case of "Cochran's Theorem" for two matrices[8].

Lemma 13 (Nabeya's Second Lemma) *Let the nonzero eigenvalues of real symmetric $n \times n$ matrices A, B, and $C = A + B$ be $\alpha_1, \dots, \alpha_a$, $\beta_1, \dots, \beta_b$ and $\gamma_1, \dots, \gamma_c$, respectively. If the relations*

$$c = a + b \tag{3.23}$$

and

$$\prod_{k=1}^{c} \gamma_k = \prod_{i=1}^{a} \alpha_i \prod_{j=1}^{b} \beta_j \tag{3.24}$$

hold, then $AB = BA = 0$.

Ogawa [162] put $s = t = 1/\lambda$ in (2.1), which then becomes, cf. (3.12) above,

$$|\lambda I_n - A| \cdot |\lambda I_n - B| = \lambda^n |\lambda I_n - A - B|;$$

this shows that the set of the nonzero eigenvalues of $A + B$ is equal to the union of the sets of the nonzero eigenvalues of A and B. So this does fulfill the conditions of Lemma 13, and hence $AB = BA = 0$ is obtained. This proof is the second complete one for the $N(0, I)$ case.

In 1949 Sakamoto [200] did not provide another proof, but gave some applications to ordinary least-squares; and he considered some χ^2 and F tests. The only inaccuracy in the paper by Sakamoto [200], regarding the Craig-Sakamoto Theorem, appears to be the author's claim that he had proved it in his 1944 paper [196]. In a footnote, Sakamoto [200] referred the reader to the 1949 paper by Matusita [138] confirming the communication they had in 1943[9].

3.2. From 1949 to 1959

3.2.1. Matérn (1949) and Uncorrelatedness

Meanwhile, the Swedish statistician Bertil Matérn [134] proved the following in 1949:

[8] Cf. e.g., Anderson and Styan (1982) [7].

[9] Sakamoto [200] wrote on page 122: "The proof of the fundamental lemma concerning the necessary condition of Theorem I [The Craig-Sakamoto Theorem] was found to be not rigorous, and I remembered that K. Matushita (sic) had given an elegant proof of the lemma for me."

Theorem 14 (Matérn 1949) *If two nonnegative quadratic forms in normally correlated variables with zero means are uncorrelated, then the two forms are independent.*

The converse of this theorem is obviously true, and hence, we can say that two nonnegative quadratic forms are independent if and only if they are uncorrelated. Furthermore, it is of interest that this result due to Matérn still holds when the dispersion matrix V is singular.

In 1950 Yukiyosi Kawada [87] extended Theorem 14 to the implication of $AB = 0$ by proving the following theorem and corollaries in 1950.

3.2.2. Kawada's 1950 Results for $\mathrm{N}(0, I)$ and "Kawada's Trace Lemma"

Theorem 15 (Kawada 1950) *Consider two quadratic forms $Q_1 = x'Ax$ and $Q_2 = x'Bx$, where A and B are real symmetric matrices not necessarily nonnegative definite and x is normally distributed with mean vector zero and dispersion matrix I. If Q_1 and Q_2 satisfy the following conditions*

$$F_{i,j} = \mathrm{E}(Q_1^i Q_2^j) - \mathrm{E}(Q_1^i)E(Q_2^j) = 0 \qquad (i, j = 1, 2), \tag{3.25}$$

then $AB = 0$.

Corollary 16 (Kawada's First Corollary) *If Q_1 and Q_2 satisfy the four conditions in equation (3.25), then Q_1 and Q_2 are independent.*

Corollary 17 (Kawada's Second Corollary) *A necessary (and sufficient) condition for the independence of Q_1 and Q_2 is $AB = 0$.*

Kawada (1950) noted that for x_k distributed $\mathrm{N}(0, 1)$,

$$\begin{aligned}
E(x_k^i) &= 0, \qquad i = 1, 3, 5, 7, \ldots \\
E(x_k^2) &= 1 \\
E(x_k^4) &= 3 \\
E(x_k^6) &= 15 \\
E(x_k^8) &= 105,
\end{aligned}$$

where $k = 1, \ldots, n$. Therefore, with $F_{(i,j)}$ defined as in (3.25), we have

$$\begin{aligned}
F_{(1,1)} &= 2\mathrm{tr}(AB) = 0 \\
F_{(1,2)} &= 8\mathrm{tr}(AB^2) + 4\mathrm{tr}(AB)\mathrm{tr}(B) = 0 \\
F_{(2,1)} &= 8\mathrm{tr}(A^2B) + 4\mathrm{tr}(AB)\mathrm{tr}(A) = 0 \\
F_{(2,2)} &= 32\mathrm{tr}(A^2B^2) + 16\mathrm{tr}((AB)^2) + 16\mathrm{tr}(AB^2)\mathrm{tr}(A) \\
&\quad + 16\mathrm{tr}(A^2B)\mathrm{tr}(B) + 8\mathrm{tr}(AB)\mathrm{tr}(A)\mathrm{tr}(B) + 8\mathrm{tr}((AB)^2) = 0.
\end{aligned}$$

From the above equations, Kawada (1950) obtained

$$\mathrm{tr}(AB) = \mathrm{tr}(AB^2) = \mathrm{tr}(AB^2) = 0$$

and hence,

$$32\mathrm{tr}(A^2B^2) + 24\mathrm{tr}((AB)^2) = 0. \tag{3.26}$$

It then follows at once from the following lemma that $AB = 0$. □

Lemma 18 (Kawada's Trace Lemma) *Let A and B be real symmetric matrices and let k be a nonnegative scalar. Then*

$$(1+k)\mathrm{tr}(A^2B^2) + \mathrm{tr}((AB)^2) \geq 0. \tag{3.27}$$

Equality holds in (3.27) *if and only if*

- $AB = -BA$ *when* $k = 0$
- $AB = 0$ *when* $k > 0$.

Proof. We expand

$$(1+k)\mathrm{tr}(A^2B^2) + \mathrm{tr}((AB)^2) = \tfrac{1}{2}\mathrm{tr}(AB+BA)'(AB+BA) + k\mathrm{tr}(AB)'AB \geq 0,$$

and our proof is complete. □

Kawada's 1950 paper [87] is the first of many important papers published in the first half of the 1950s. In 1950 alone, these four publications appeared:

- Aitken (1950) [3]
- Carpenter (1950) [27]
- Kawada (1950) [87]
- Ogawa (1950) [164].

Carpenter's article [27] contains the conditions for a quadratic form to be distributed as chi-square and conditions for the independence of two quadratic forms in the $N(\mu, I)$ case. This appears to be the first treatment of the noncentral case where the mean μ is not necessarily 0.

Carpenter's proof for the Craig-Sakamoto Theorem goes as follows. First, he presented the moment generating functions $G(s,t)$, $G(s,0)$, and $G(0,t)$ of $x'(A+B)x$, $x'Ax$, and $x'Bx$, respectively; and he showed that $G(s,t) = G(t,0)\cdot G(0,s)$ holds when $AB = 0$. Unfortunately, to demonstrate the necessity part, he referred to Hotelling (1944) and Craig (1943), which as we have already noted, are both unsatisfactory proofs. Moreover he did not establish the necessary Laha's Lemma, cf. Lemma 20 below. However, Aitken [3] and Ogawa [162] both presented correct proofs. Aitken's 1950 proof in [3] is similar to Ogawa's 1949 proof in [162] published a year earlier.

3.2.3. Aitken's 1950 Proof for $\mathrm{N}(0, V)$, with V Positive Definite

Assuming independence, Aitken in 1950 noticed [3] that

$$|\lambda I_n - A| \cdot |\lambda I_n - B| = \lambda^n |\lambda I_n - A - B|, \tag{3.28}$$

cf. (3.12), but unlike Ogawa (1946), Aitken (1950) obtained (3.28) not by putting $s = t = 1/\lambda$ in (2.1), but because

> " ... the latent roots of the matrix pencil $sA + tB$ are the latent roots of sA together with those of tB, with the useful corollary that the rank of $sA + tB$ is the sum of the ranks of sA and tB [for all real s and t]."

To prove his claim, he diagonalized A into A_{11} and without lost of generality, he applied an orthogonal transformation to $sA + tB$ and showed that if

$$\begin{vmatrix} sA_{11} + tB_{11} & tB_{12} \\ tB'_{12} & tB_{22} \end{vmatrix} = \begin{vmatrix} sA_{11} & . \\ . & tB_{22} \end{vmatrix} \tag{3.29}$$

then $B_{12} = 0$, and the submatrices A_{11} and B_{22}, which are the upper left and the lower right submatrix, respectively, of the diagonalized A and B, are disjoint. Therefore $AB = 0$.

In this same paper, Aitken (1950) introduced the transformation $y = V^{1/2}x$ that produced uncorrelated variates and referring to the proof he gave for the $\mathrm{N}(0, I)$ case, he extended the result to $AVB = 0$ for the $\mathrm{N}(0, V)$ case with V being positive definite. We observe that though Aitken was almost certainly not aware of the Japanese work on this topic, his ideas and results are similar to those proposed by Sakamoto (1944) and the proof of the Craig-Sakamoto Theorem given by Ogawa (1949) using Nabeya's lemmas.

3.2.4. Ogawa's 1950 Proof for $\mathrm{N}(\mu, I)$

The 1950 paper by Ogawa [164] contains the first correct proof for the noncentral case $\mathrm{N}(\mu, I)$, and is, therefore, the most important development since the first complete proof of the Craig-Sakamoto Theorem by Matusita (in 1949 in [138]). Ogawa [164] proved the following:

Theorem 19 (The Craig-Sakamoto Theorem with $x \sim \mathrm{N}(\mu, I)$) *Let A and B be symmetric matrices and $x \sim \mathrm{N}(\mu, I)$ with mean vector μ and dispersion matrix I. Then $x'Ax$ and $x'Bx$ are independent if and only if $AB = 0$.*

Unfortunately, as for Matusita (1949), Ogawa's 1950 paper did not receive the recognition it should have had at the time. After proving sufficiency, Ogawa converted the assumption of independence into an equation involving the moment generating functions of the quadratic forms $x'Ax, x'Bx$, and $x'(A + B)x$ to obtain a ratio of characteristic equations that is equal to the exponent of a ratio of polynomials. Then, he stated—and proved—the following lemma, which is the version for real variables of what is now known as "Laha's Lemma", cf. Lemma 22 below.

Lemma 20 (Laha's Lemma for Real Variables) *If $P(t)$, $Q(t)$, $R(t)$, and $S(t)$ are real polynomials in t and*

$$\exp\{P(t)/Q(t)\} = R(t)/S(t) \tag{3.30}$$

then both $P(t)/Q(t)$ and $R(t)/S(t)$ are constants.

Hence, Ogawa (1950) proved the necessity of the Craig-Sakamoto Theorem. We notice that if we reduce the $\mathrm{N}(0, V)$ case to $\mathrm{N}(0, I)$—as proposed by Sakamoto (1944) and Aitken (1950)—then we obtain a proof for the general case of the Craig-Sakamoto Theorem. Ogawa's proof is included below as Ogawa and Laha's proof (subsection 3.2.7) for the general case with V being positive definite.

In 1951, Ogasawara and Takahashi published a paper [157] on the independence of quadratic forms and the conditions they must satisfy to obtain a chi-squared distribution. This 1951 paper [157] contains a nice and simple proof of the $\mathrm{N}(0, I)$ case using determinants, traces, and series expansions. Moreover, it is the first paper to establish (completely and correctly) necessary and sufficient conditions for independence of two quadratic forms when the underlying dispersion matrix is not necessarily nonsingular.

3.2.5. Ogasawara and Takahashi's 1951 Proof for $\mathrm{N}(0, I)$

Ogasawara and Takahashi [157] showed that:

$$AB = 0 \quad \Leftrightarrow \quad |I - sA - tB| = |I - sA| \cdot |I - tB| \qquad \forall \text{ real } s \text{ and } t. \tag{3.31}$$

Hence, by Theorem 9, it follows from the above statement that $x'Ax$ and $x'Bx$ are independent. As we have already noted, it is obvious that the first equation implies the second. To show the converse we use Lemma 7 in Chapter 1. Assuming that

$$|I - sA - tB| = |I - sA| \cdot |I - tB| \qquad \forall \text{ real } s \text{ and } t \tag{3.32}$$

holds, Ogasawara and Takahashi [157] multiplied both sides of the (3.32) by $|K|$, where $K = (I - sA)^{-1}$, and obtained:

$$\begin{aligned} |I - sA|^{-1} \cdot |I - sA| \cdot |I - tB| &= |I - sA|^{-1} \cdot |I - sA - tB| \\ |I - tB| &= |I - tKB|. \end{aligned} \tag{3.33}$$

Taking logarithms and using Lemma 7, gives

$$\sum_{j=1}^{\infty} \frac{t^j}{j} \mathrm{tr}(B^j) = \sum_{j=1}^{\infty} \frac{t^j}{j} \mathrm{tr}(KB)^j \tag{3.34}$$

for all real (s, t) sufficiently near $(0, 0)$. We notice that all the crossproducts in s and t on the right-hand side must be zero. Setting $j = 4$ and after some algebraic manipulation, we find that the coefficient of s^2t^2 is

$$4\mathrm{tr}(A^2B^2) + 2\mathrm{tr}((AB)^2) = 0, \tag{3.35}$$

and so $AB = 0$ follows at once from Kawada's Trace Lemma (our Lemma 18 above).

Ogasawara and Takahashi [157] extended the above result to cover the case where V is non-negative definite. This appears to be first complete correct treatment for this most general case of $\mathrm{N}(\mu, V)$, with V possibly singular.

3.2.6. Ogasawara and Takahashi's 1951 Proof with V Possibly Singular

Theorem 21 (The Craig-Sakamoto Theorem for the $\mathrm{N}(\mu, V)$ case) *Let $x \sim \mathrm{N}(\mu, V)$ with V possibly singular, and let A, B be symmetric matrices. Then $x'Ax$ and $x'Bx$ are independent if and only if the following conditions hold*

$$\begin{aligned} VAVBV &= 0 && (3.36)\\ VAVB\mu &= 0 && (3.37)\\ VBVA\mu &= 0 && (3.38)\\ \mu'AVB\mu &= 0. && (3.39) \end{aligned}$$

Ogasawara and Takahashi [157] used the "symmetrized form" of the moment-generating function of $x'Ax$, cf. Theorem 1[10] above:

$$\mathrm{E}(\exp\{sx'Ax\}) = \frac{\exp\{\mu'(sA + 2t^2AU(I - 2sUAU)^{-1}UA)\mu)\}}{|I - 2sUAU|^{1/2}}, \tag{3.40}$$

where $U = U'$ and $U^2 = V$.

Assuming independence, equating the product of the moment generating functions of $x'Ax$ and $x'Bx$ with the moment generating function of $x'(A+B)x$ yields, using Laha's Lemma and after some algebraic manipulations, for all real s and t:

$$|I - 2sA_1| \cdot |I - 2tB_1| = |I - 2sA_1 - 2tB_1| \tag{3.41}$$

and

$$\begin{aligned} &\mu'(s^2AU(I - 2sA_1)^{-1}UA + t^2BU(I - 2tB_1)^{-1}UB)\mu \\ &= \mu'((sA + tB)U(I - 2sA_1 - 2tB_1)^{-1}U(sA + tB))\mu, \end{aligned} \tag{3.42}$$

where $A_1 = UAU$ and $B_1 = UBU$. Ogasawara and Takahashi [157] then showed that (3.41) implies $A_1B_1 = 0$ and thus $VAVBV = 0$, i.e., (3.36), in a way similar to Kawada (1950), cf. Kawada's Trace Lemma (our Lemma 18 above).

We now use Lemma 7 and substitute the geometric series expansions for the inverses in (3.42) to obtain:

$$\begin{aligned} &\mu'\left\{s^2AU\left(\sum_{h=0}^{\infty}(2sA_1)^h\right)UA + t^2BU\left(\sum_{k=0}^{\infty}(2sB_1)^k\right)UB\right\}\mu \\ &= \mu'\left\{((sA + tB)U\left(\sum_{\ell=0}^{\infty}(2sA_1 + 2tB_1)^{\ell}\right)U(sA + tB)\right\}\mu. \end{aligned} \tag{3.43}$$

[10] Theorem 1 gives the characteristic function of $x'Ax$ which is equal to the moment-generating function of $ix'Ax$.

Equating the coefficient of st equal to 0 yields

$$\mu' AVB\mu + \mu' BVA\mu = 0.$$

But $\mu' AVB\mu$ is a scalar and so equals $(\mu' AVB\mu)' = \mu' BVA\mu$. Thus $\mu' AVB\mu = 0$ as in (3.39).

Equating the coefficient of s^2t^2 equal to 0 yields

$$\mu' AVAVBVB\mu + \mu' AVBVAVB\mu + \mu' AVBVBVA\mu + \mu' BVAVAVB\mu \\ + \mu' BVAVBVA\mu + \mu' BVBVAVA\mu = 0. \quad (3.44)$$

Substituting $VAVBV = 0 = VBVAV$ from (3.15) yields the two middle terms of (3.44)

$$\mu' AVBVBVA\mu + \mu' BVAVAVB\mu = 0.$$

But

$$\mu' AVBVBVA\mu + \mu' BVAVAVB\mu = (AVB\mu)'V(AVB\mu) + (BVA\mu)'V(BVA\mu) \geq 0 \quad (3.45)$$

with equality if and only if (3.37) and (3.38) hold. Our proof is complete. □

Ogasawara and Takahashi's paper [157] appears to be the last one written by Japanese researchers on the Craig-Sakamoto Theorem until the 1980s. It would take more than thirty years before the publication in Japan of a generalization of Ogasawara and Takahashi's results to the Wishart distribution by Hyakutake and Siotani (1985).

Meanwhile, researchers in the United States, who were unaware of the progress made in Japan, were still trying to find a complete proof of the characterization they called Craig's Theorem. In 1951, Nelder [153] and in 1952, Lukacs [122] published papers related to the independence of quadratic forms, but they did not explicitly cover the Craig-Sakamoto Theorem. However, in 1954, Lancaster [113] presented a proof for the simple case $N(0, I)$ that is similar to Aitken (1950) and to Matusita (1949), but Lancaster used traces and cumulants. Furthermore, in his historical account Lancaster [113] mentioned that Hotelling (1944) had questioned the validity of the theorem and that Ogawa (1949) had drawn attention to the lacuna in Hotelling's (1944) proof; Lancaster [113] also commented on the results by Matérn (1949) and Kawada (1950) on the relation between the covariance between and the independence of two quadratic forms.

3.2.7. Proof by Laha (1956) for $N(\mu, I)$ and "Laha's Lemma"

In 1956 Laha [110] introduced a different proof of the Craig-Sakamoto Theorem for the $N(\mu, I)$ case, and he extended the result to second-degree polynomials (or bilinear forms). In his proof, he stated the following lemma:

Lemma 22 (Laha's Lemma for Complex Variables) *If the relation*

$$\exp[P(it_1, it_2)/Q(it_1, it_2)] = R(it_1, it_2)/S(it_1, it_2) \quad (3.46)$$

holds for all real t_1 and t_2, where $i = \sqrt{-1}$ and P, Q, R, and S are polynomials in t_1 and t_2, then the rational functions are constants.

The proof then follows easily. Unfortunately, Laha omitted a proof for this Lemma 22. Later Searle [208], Gundberg [62], Driscoll and Gundberg [42], among others, became interested in this result which has been referred to by many as "Laha's Lemma". However, a similar proof had already been published in 1950 by Ogawa [164], which few had noticed at the time. Moreover, in 1956 Laha [110] omitted the facts that the Craig-Sakamoto Theorem was not correctly proved by Craig (1943), nor by Sakamoto (1944), nor by Hotelling (1944). Nevertheless, Laha did refer to the papers of Matusita (1949) and Ogawa (1949) for a correct proof, and Laha's proof of the Craig-Sakamoto Theorem, providing that Laha's Lemma holds, was nice and attracted attention.

Theorem 23 (The Craig-Sakamoto Theorem with $x \sim N(\mu, V)$) *Let A and B be symmetric matrices and $x \sim N(\mu, V)$ with mean vector μ and positive definite dispersion matrix V, then $x'Ax$ and $x'Bx$ are independent if and only if $AVB = 0$.*

Using the transformation $x \to V^{1/2}y$, one reduces this case to the $N(\mu, I)$ case. We know by Theorem 1 that the moment generating function of $x'Ax$ can be expressed in the following way:

$$G(s,0) = \mathrm{E}(\exp\{sx'Ax\}) = \frac{\exp[s\mu'(I-2sAV)^{-1}A\mu]}{|I-2sAV|^{1/2}} \tag{3.47}$$

which reduces to

$$\phi(s) = \frac{\exp\{s\mu'(I-2sA)^{-1}A\mu]\}}{|I-2sA|^{1/2}}$$

when I is the dispersion matrix. After some calculations using the moment generation functions $G(s,0)$, $G(0,t)$ and $G(s,t)$ as given in Theorem 1 with $V = I$, we have

$$\left(\frac{|I-2sA|^{1/2}|I-2tB|^{1/2}}{|I-2sA-2tB|^{1/2}}\right)^{\frac{1}{2}} =$$
$$\exp\{\mu'[s(I-2sA)^{-1}A + t(I-2tB)^{-1}B - st(I-2sA-2tB)^{-1}(sA+tB)]\mu\} \tag{3.48}$$

if and only if $G(s,0)\cdot G(0,t) = G(s,t)$.

Proof of Sufficiency. Assume that $AB = 0$, then we obtain

$$|I-2sA||I-2tB| = |I-2sA-2tB| \quad \Rightarrow \quad \frac{|I-2sA||I-2tB|}{|I-2sA-2tB|} = 1. \tag{3.49}$$

Now, since the left-hand side of equation (3.48) is equal to 1, it follows that the expression in the exponent on the right-hand side is zero. So we have

$$\frac{\mu'(sA+tB)\mu}{|I-2sA|^{1/2}|I-2tB|^{1/2}} = \frac{\mu'(st(sA+tB))\mu}{|I-2sA-2tB|^{1/2}}. \tag{3.50}$$

And hence, since the numerators and the denominators of equation (3.48) are equal, equality holds and independence is shown. □

Proof of Necessity. Assuming independence and letting the numerator and the denominator of the left-hand side of (3.48) equal $R(t)$ and $S(t)$, respectively, and the numerator and the denominator of the right-hand side of (3.48) equal $P(t)$ and $Q(t)$, respectively, we obtain that $\exp(P(t)/Q(t)) = R(t)/S(t)$ is constant.

This result alone became famous, since Laha's 1956 proof of the Craig-Sakamoto Theorem did not include a proof of Laha's Lemma (our Lemma 22). At the time, few were aware of the research published in Japan and it was thought that Laha's Lemma, which involved complex analysis, was the only way to prove the Craig-Sakamoto Theorem. Fortunately, a function-theoretic proof in the complex plane was already given by Ogawa [164] in 1950. Ogawa [164] assumed that $|s| < \varepsilon$ and $|t| < \varepsilon$, he fixed s so that the polynomials P, R, S, and T are only in t, and he proved by contradiction that $P(t)/Q(t)$ had to be constant.

We now rephrase Laha's Lemma (our Lemma 22 above) as follows:

Lemma 24 (Laha's Lemma for Complex Variables: Alternate Version) *Suppose that q and r are rational functions (quotients of polynomials) and that*

$$\exp\{q(u)\} = r(u)$$

for all values of the real scalar u in some non-empty region in $\mathbb{R}$. Then q and r are constant.

Proof[11]. We use the standard classification of isolated singularities of analytic functions. There are three kinds of such singularites, cf. e.g., Rudin [195, pp. 210–211]:

(1) Removable singularities—singularities that are really not there at all

(2) Poles

(3) Essential singularities.

Rational functions have only removable singularities and poles. The other key facts are:

(a) If $q(u)$ has a removable singularity at z then $\exp\{q(u)\}$ also has a removable singularity there.

(b) If $q(u)$ has a pole at z then $\exp\{q(u)\}$ has an essential singularity at z.

So, by analytic continuation, we extend q and r to the Riemann sphere minus the finite set of (necessarily) isolated singularities. The relation $e^{q(u)} = r(u)$ continues to hold on this set. But $r(u)$ has only removable singularities and poles and $e^{q(u)}$ has only removable singularities and essential singularities. Therefore, all singularities are removable. But a rational function with only removable singularities on the Riemann sphere is necessarily constant. Hence q and r are constant. Our proof is complete. □

Hence, $R(t)/S(t)$ must be constant and equal to 1 by assumption of independence and this equality implies that $|I-2sA||I-2tB| = |I-2sA-2tB|$ in the $N(\mu, I)$ case, which in turn implies that $AB = 0$, and in the $N(\mu, V)$ case using the transformation $V^{-1/2}y \rightarrow x$ with V positive definite, we obtain $AVB = 0$.

[11] I am very grateful to Professor S. W. Drury (McGill University) for giving me this proof.

3.2.8. Proof by Taussky (1958) using Matrices with Property L

Another important paper on the Craig-Sakamoto Theorem, and well worth attention, is the 1958 paper by Olga Taussky [231] . She explained very clearly the conditions in which the Craig-Sakamoto Theorem holds, using matrices with the so-called property L, which was introduced in 1952 by Motzkin and Taussky [148]. The square complex matrices A and B are said to have property L whenever the eigenvalues of $sA+tB$ are equal to $s\alpha_j + t\beta_j$ for all values of s and t and for a certain fixed pairing of the eigenvalues α_j of A and β_j of B. Then Taussky [231] noted (in paragraph γ on page 139) that "A pair of hermitian matrices with property L is commutative," referring to the 1952 paper by Motzkin and Taussky [148] for a proof. Moreover Taussky goes on to note that "This is even true for normal matrices." Taussky [231] proved that if the real symmetric matrices A and B satisfy

$$|\lambda I_n - A| \cdot |\lambda I_n - B| = \lambda^n |\lambda I_n - A - B|,, \tag{3.51}$$

cf. (3.28), then A and B have property L.

See also the proof discussed above by Aitken [3] and the new proof in Theorem 32 below by Li and Styan [118] in our Chapter 4.

3.3. From 1960 to 1979

After the publication of the proof by Ogawa (1950) and Laha (1956), efforts were made to find further results and correct proofs of the Craig-Sakamoto Theorem. In 1960, Laha and Lukacs [111] extented Kawada's (1950) result to the following:

Theorem 25 (Laha and Lukacs 1960) *Let $x \sim N(0, I)$. If $Q = x'Ax + b'x$ and $L = c'x$ are uncorrelated of order* $(2, 2)$, *that is* $\mathrm{E}(Q^i L^j) = \mathrm{E}(Q^i)\mathrm{E}(L^j)$ *for $i = 1, 2$ and $j = 1, 2$, then $c'A = 0$ and $c'b = 0$.*

From this theorem, they then deduced that Q and L are independent if and only if they are uncorrelated of order $(2, 2)$.

3.3.1. Pao-Lu Hsu's 1962 Lecture

According to Fang and Zhang (1990) [52, Lemma 2.8.3, pp. 77–79][12] Pao-Lu Hsu presented in a 1962 lecture[13] [74] the following theorem[14]:

Theorem 26 (Pao-Lu Hsu 1962) *Let A and B be real symmetric matrices with $\alpha_1, ..., \alpha_a$ and $\beta_1, ..., \beta_b$ being the nonzero eigenvalues of A and B respectively, with $a = \mathrm{rank}(A)$ and $b = \mathrm{rank}(B)$. If the nonzero eigenvalues of $A + B$ are $\alpha_1, ..., \alpha_a, \beta_1, ..., \beta_b$, then $AB = BA = 0$.*

[12] See also Zhang and Fang [247, Lemma 2.8.3, pp. 123–125].

[13] Apparently not published by Pao-Lu Hsu.

[14] I am very grateful to Chang-Yu Lu and Bao-Xue Zhang for bringing this to my attention, and to Ka Lok Chu and Professor Kai-Tang Fang (Hong Kong Baptist University) for providing me with a copy of Zhang and Fang [247].

Proof. First, assume that $a+b=n$ so that $A+B$ is of full rank. Since we can diagonalize A, then without loss of generality, let

$$A=\begin{pmatrix} D_a & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.52}$$

where $D_a = \text{diag}(\alpha_1, ..., \alpha_a)$. Since B is symmetric, there exists an orthogonal matrix P such that

$$P'BP=\begin{pmatrix} 0 & 0 \\ 0 & D_b \end{pmatrix}, \tag{3.53}$$

where $D_b = \text{diag}(\beta_1, ..., \beta_b)$. We now partition the matrix P similarly to A and $P'BP$,

$$P=\begin{pmatrix} C & F \\ E & G \end{pmatrix}, \tag{3.54}$$

with C is $a \times a$ and G is $b \times b$. Then

$$B=\begin{pmatrix} C & F \\ E & G \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & D_b \end{pmatrix}\begin{pmatrix} C' & E' \\ F' & G' \end{pmatrix}=\begin{pmatrix} I & F \\ 0 & G \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & D_b \end{pmatrix}\begin{pmatrix} I & 0 \\ F' & G' \end{pmatrix}. \tag{3.55}$$

It is easy to see that,

$$A=\begin{pmatrix} I & F \\ 0 & G \end{pmatrix}\begin{pmatrix} D_a & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I & 0 \\ F' & G' \end{pmatrix}. \tag{3.56}$$

Therefore,

$$A+B=\begin{pmatrix} I & F \\ 0 & G \end{pmatrix}\begin{pmatrix} D_a & 0 \\ 0 & D_b \end{pmatrix}\begin{pmatrix} I & 0 \\ F' & G' \end{pmatrix}. \tag{3.57}$$

Since the set of nonzero eigenvalues of XY is the same as the set of nonzero eigenvalues of YX for any pair of conformable matrices X and Y, it follows that the set of nonzero eigenvalues of $A+B$ equals the set of nonzero eigenvalues of

$$H=\begin{pmatrix} D_a & 0 \\ 0 & D_b \end{pmatrix}\begin{pmatrix} I & 0 \\ F' & G' \end{pmatrix}\begin{pmatrix} I & F \\ 0 & G \end{pmatrix}=\begin{pmatrix} D_m & 0 \\ 0 & D_n \end{pmatrix}\begin{pmatrix} I & F \\ F' & F'F+G'G \end{pmatrix}. \tag{3.58}$$

Since P is orthogonal we have $F'F+G'G=I$. So,

$$H=\begin{pmatrix} D_a & 0 \\ 0 & D_b \end{pmatrix}\begin{pmatrix} I & F \\ F' & I \end{pmatrix}. \tag{3.59}$$

Since these determinants are equal:

$$\begin{vmatrix} I & F \\ F' & I \end{vmatrix} = |I - F'F|, \tag{3.60}$$

we obtain

$$\begin{aligned} \prod_{i=1}^{r} \alpha_i \prod_{j=1}^{s} \beta_j = |A+B| = |H| = \begin{vmatrix} D_a & 0 \\ 0 & D_b \end{vmatrix} \begin{vmatrix} I & F \\ F' & I \end{vmatrix} &= \prod_{i=1}^{a} \alpha_i \prod_{j=1}^{b} \beta_j |I - F'F|, \\ &\Rightarrow |I - F'F| = 1 \\ &\Rightarrow F'F = 0 \Rightarrow F = 0. \end{aligned} \tag{3.61}$$

The implication (3.61) follows from the fact that the eigenvalues of $F'F$ are all nonegative and at most equal to 1, since $F'F + G'G = I$. Hence $G'G = I$ and so

$$B = \begin{pmatrix} I & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & D_b \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & G' \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & GD_bG' \end{pmatrix} \tag{3.62}$$

and $AB = BA = 0$. Moreover, it can easily be proven that the theorem holds in the case where $r + s < n$ by a similar argument. □

In the same year, Bhat (1962) [21] showed that a quadratic form Q is independent of the sum of a finite number of nonnegative definite quadratic forms $Q_1 + Q_2 + \cdots + Q_n$ if it is independent of each of them separately. During the years 1962 and 1963 in India, the late C. G. Khatri in [93] and [94] established necessary and sufficient conditions for second-degree polynomials in normal vectors to be independently distributed or to follow Wishart distributions. Moreover, he extended Bhat's results and showed that Rao's result (1962) on quadratic forms, when B is singular, is false if it is not specified that the degrees of freedom of the chi-square distribution were equal to the rank of V. As for the conditions he gave for the independence of second-degree matrix polynomials, they are equivalent to the ones found by Ogasawara and Takahashi (1951) for quadratic forms[15].

In the following year, Good (1963) [56] presented a series of results that immediately follow from the Craig-Sakamoto Theorem and that cover both quadratic and linear forms. His first Theorems 1 and 1C are equivalent to the case where $x \sim N(0, V)$ with V being non-singular in Theorem 1 and V being possibly singular in Theorem 1C. The conditions Good found are the same as stated by Sakamoto [196], Matusita [138], and Ogasawara and Takahashi [157]. In addition, Good presented two other theorems (his Theorems 2 and 3) on the independence of $x'Ax + a'x$ and $x'Bx + b'x$ and on the independence of more than two quadratic forms. Although Good (1963) contains some lacunas that were noticed by Shanbhag (1966) [218], these were easily corrected by Good (1966) [56]. For his part, Shanbhag (1966) showed that for A being nonnegative definite, the conditions for independence are reduced to $AVBV = 0$ and $AVB\mu = 0$; and if both A and B are nonnegative definite then $AVB = 0$ is necessary and sufficient for independence.

The Craig-Sakamoto Theorem was gaining in popularity and appeared in four books:

[15]See also Styan (1970) [225]

- *The Advanced Theory of Statistics, Vol. 1: Distribution Theory*, by Kendall and Stuart [89];
- *Continuous Univariate Distributions*, by Johnson and Kotz [83];
- *Introduction to Mathematical Statistics*, by Hogg and Craig [72]; and
- *Linear Models*, by Searle [209].

However, many found that the information available in these books was insufficient.

3.3.2. Searle's *Linear Models* Proof for $\mathrm{N}(\mu, V)$

As noticed by Nagase and Banerjee (1973) [150], Searle (1984) [208], Provost (1994) [175] and others, the proof given in the well-known *Linear Models* book by Searle [209, Th. 4, pp. 59–60] is incomplete. Searle stated that with $x \sim \mathrm{N}(\mu, V)$, with V possibly singular, then $AVB = 0$ if and only if $x'Ax$ and $x'Bx$ are independent. To prove this, he proceeded as follows: Independence of $x'Ax$ and $x'Bx$ implies that their covariance is equal to zero. Hence,

$$\mathrm{var}(x'Ax + x'Bx) - \mathrm{var}(x'Ax) - \mathrm{var}(x'Bx) = 0$$

$$2\mathrm{tr}[(A+B)V]^2 + 4\mu'(A+B)V(A+B)\mu - 2\mathrm{tr}(AV) - 4\mu'AVA\mu - 2\mathrm{tr}(BV)^2 - 4\mu'BVB\mu = 0$$

and

$$\mathrm{tr}(AVBV) + 2\mu'AVB\mu = 0. \tag{3.63}$$

So far, the argument is correct. Searle, however, then claimed that (3.63) holds <u>for all</u> μ and therefore

$$\mathrm{tr}(AVBV) = 0; \tag{3.64}$$

but (3.63) holds <u>only for the specified</u> $\mu = \mathrm{E}(x)$, and (3.64) alone does not imply

$$\begin{pmatrix} V \\ \mu \end{pmatrix} AVB(V : \mu) = 0. \tag{3.65}$$

Searle also claimed that

$$\mathrm{tr}(AVBV) = 0 \Rightarrow AVB = 0; \tag{3.66}$$

but this implication does not hold in general for if $A = V = I$ and

$$B = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{3.67}$$

then $\mathrm{tr}(AVBV) = \mathrm{tr}(B) = 0$ but $AVB = B \neq 0$.

When A and B are both nonnegative definite, however, the implication does hold—we may now write $A = SS'$ and $B = TT'$ and

$$\mathrm{tr}(AVBV) = \mathrm{tr}(SS'VTT'V) = \mathrm{tr}(S'VTT'VS) = \mathrm{tr}[S'VT(S'VT)'] \geq 0 \tag{3.68}$$

with equality if and only if $S'VT = 0 \Leftrightarrow AVB = 0$.

In the early 1970s, several researchers tried to clarify various aspects of the development of the Craig-Sakamoto Theorem. Scarowsky's MSc thesis (1973) [203] gave a good review of the conditions required for independence for quadratic, linear, and bilinear forms. His thesis is a reliable source of correct proofs and the history of the theorem. It contains one of the most complete lists of references available, and the bibliography in the present thesis was built on it.

Two extensions to Laha's result (1956) [110] (one in 1973 and the other in 1976) and a counter-example to Searle's proof were presented by Nagase and Banerjee [150, 152]. A nice historical account of the Craig-Sakamoto Theorem was presented in a 1974 lecture by Rayner [189]. Moreover, a 1977 paper by Tan [230] filled in more details, and in 1978 Krafft [108] presented a different proof.

3.3.3. Krafft's 1978 Proof

Following the proof by Searle in his *Linear Models* book as just discussed, in 1978 Krafft [108] started by noting that independence implies that the covariance $\mathrm{cov}(\mathbf{x}'A\mathbf{x}, \mathbf{x}'B\mathbf{x}) = 0$ and so he obtained

$$\mathrm{tr}(AVBV) + 2\mu'AVB\mu = 0. \tag{3.69}$$

Krafft [108] then claimed that (3.69) implies that $\mu'AVB\mu = 0$ for all μ. Setting $u + v = \mu$ we obtain

$$(u' + v')AVB(u + v) = 0$$

$$u'AVBu + v'AVBv + u'AVBv + v'AVBu = 0. \tag{3.70}$$

Thus $u'AVBu = 0$ and $v'AVBv = 0$; and knowing that $(AVB)' = BVA$, we obtain

$$u'(AVB + BVA)v = 0 \qquad \forall \quad u, v. \tag{3.71}$$

Letting u and v be, in turn, the columns of an identity matrix, (3.71) yields

$$AVB + BVA = 0 \Rightarrow AVB = -BVA. \tag{3.72}$$

To show that this implies $AVB = 0$, Krafft [108] now put $s = t$ in the equation

$$|I - sA| \cdot |I - tB| = |I - sA - tB|, \tag{3.73}$$

which must hold under the assumption of independence. Then, letting $TT' = V$, $T'AT = \dot{A}$ and $T'BT = \dot{B}$, equation (3.73) becomes

$$|I - t\dot{A}| \cdot |I - t\dot{B}| = |I - t(\dot{A} + \dot{B})|. \tag{3.74}$$

Since $\dot{A}$ and $\dot{B}$ are symmetric and the numbers of their nonzero eigenvalues are equal to their ranks, the determinants on the left-hand side of (3.74) are polynomials in t, of order $\mathrm{rank}(A)$ for the first term and rank(B) for the second; and the determinant on the right-hand side is a polynomial of order rank($A+B$). Thus we have rank additivity:

$$\text{rank}(A) + \text{rank}(B) = \text{rank}(A+B). \tag{3.75}$$

Diagonalizing $\dot{A}$ with the orthogonal matrix U yields

$$U'\dot{A}U = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.76}$$

where the diagonal matrix D is $a \times a$ and nonsingular, with $a = \text{rank}(\dot{A}) = \text{rank}(A)$. We write

$$U'\dot{B}U = G = \begin{pmatrix} G_{11} & G_{12} \\ G'_{12} & G_{22} \end{pmatrix}, \tag{3.77}$$

where G_{11} is $a \times a$. Multiplying these two matrices yields

$$U'\dot{A}UU'\dot{B}U = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} G_{11} & G_{12} \\ G'_{12} & G_{22} \end{pmatrix}. \tag{3.78}$$

Since U is orthogonal and $TT' = V$, $T'AT = \dot{A}$, $T'BT = \dot{B}$, this becomes

$$U'T'AVBTU = \begin{pmatrix} DG_{11} & DG_{12} \\ 0 & 0 \end{pmatrix}. \tag{3.79}$$

Similarly for $-BVA$,

$$-U'\dot{B}UU'\dot{A}U = -U'T'BVATU = -\begin{pmatrix} G_{11} & G_{12} \\ G'_{12} & G_{22} \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} G_{11}D & 0 \\ G'_{12}D & 0 \end{pmatrix}. \tag{3.80}$$

Since $AVB = -BVA$, we have

$$\begin{pmatrix} DG_{11} & DG_{12} \\ 0 & 0 \end{pmatrix} = -\begin{pmatrix} G_{11}D & 0 \\ G'_{12}D & 0 \end{pmatrix}. \tag{3.81}$$

For this equality to be true, G_{12} must be the zero matrix. Thus

$$U'T'AVBTU = \begin{pmatrix} DG_{11} & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.82}$$

and so

$$U'\dot{B}U = \begin{pmatrix} G_{11} & 0 \\ 0 & G_{22} \end{pmatrix}. \tag{3.83}$$

Adding $U'\dot{A}U$ and $U'\dot{B}U$, we have

$$U'\dot{A}U + U'\dot{B}U = \begin{pmatrix} D+G_{11} & 0 \\ 0 & G_{22} \end{pmatrix}. \tag{3.84}$$

Now, one can see that $\text{rank}(A+B) = \text{rank}(D+G_{11}) + \text{rank}(G_{22}) = \text{rank}(A) + \text{rank}(B)$ and that diagonalizing $\dot{A}$ implies $\text{rank}(A) = \text{rank}(D)$, so that $\text{rank}(B) = \text{rank}(G_{11}) + \text{rank}(G_{22})$, giving

$$\text{rank}(D+G_{11}) = \text{rank}(D) + \text{rank}(G_{11}) \geq \text{rank}(D) = a, \tag{3.85}$$

since the $a \times a$ matrix D has full rank and $\text{rank}(G_{11}) \geq 0$. Hence $G_{11} = 0$ and thus

$$U'\dot{B}U = \begin{pmatrix} 0 & 0 \\ 0 & G_{22} \end{pmatrix}. \tag{3.86}$$

Therefore, $U'T'AVBTU = 0$ and $AVB = 0$.

While this proof by Krafft [108] does correctly show that

$$AVB = -BVA \quad \text{and} \quad |I-tA|\cdot|I-tB| = |I-tA-tB| \quad \Rightarrow \quad AVB = 0, \tag{3.87}$$

the assertion that independence of $x'Ax$ and $x'Bx$ implies that

$$\text{tr}(AVBV) + 2\mu'AVB\mu = 0 \tag{3.88}$$

must hold for all μ is clearly false, cf. Searle's proof above[16].

3.4. From 1980 to 1996

During the eighties and early nineties, research on the Craig-Sakamoto Theorem focused on distinguishing the correct proofs from the unsatisfactory ones, on completing the incomplete ones, and bringing to light some proofs that were considered to deserve more recognition.

The 1985 paper written by Zielinski [248] mentions a very short proof and may leave the reader confused because of its lack of explanations. One should understand that Zielinski's goal was to shorten Ogawa's 1949 [162] proof for the $N(0, I)$ case which, in this article [248], he called "Nabeya's" proof. Zielinski claimed that "his proof was shorter than the one presented by Rao and Mitra (1971) [185] and that there was no need to rely on convergence in Banach spaces".

3.4.1. Zielinski's 1985 Proof

Using Lemma 13, in 1985 Zielinski [248] let $a_1, ..., a_r$; $b_1, ..., b_s$; $g_1, ..., g_t$ be orthonormal systems of eigenvectors of A, B, and C, respectively, and each of these vectors are related to the eigenvalues $\alpha_1, ..., \alpha_r$; $\beta_1, ..., \beta_s$; $\gamma_1, ..., \gamma_t$ of eigenvectors of A, B, and C. Since $C = A + B$, then the vectors

[16] See subsection 3.3.2 of this Chapter.

$a_1, \ldots, a_r$; $b_1, \ldots, b_s$ form a basis for $\mathcal{R}(C)$. Then, Zielinski [248] stated that by comparison of the matrix representations of the linear operator $A + B : \mathcal{R}(C) \to \mathcal{R}(C)$ at the given bases, the corresponding Gramian determinants coincide. However, as Ogawa (1993) [165] showed, this is only true when $|L'L| = 1$, where L and L' are the matrices that diagonalize A and B. Therefore, the vectors $a_1, \ldots, a_r$; $b_1, \ldots, b_s$ are orthogonal. In others words, Zielinski claimed that this argument is necessary and sufficient to show that for each c_k there corresponds a unique a_i or b_j for all i, j, k or that this correspondence is isomorphic. Therefore, $\alpha_1 \cdot \ldots \cdot \alpha_r \cdot \beta_1 \cdot \ldots \cdot \beta_s = \gamma_1 \cdot \ldots \cdot \gamma_t$ and $t = s + r$ which implies $AB = 0$.

The next proof was first published in 1988 by Reid and Driscoll [192] and revised in 1995 by Driscoll and Krasnicka [43].

3.4.2. Surveys by Reid & Driscoll [192] and Driscoll & Krasnicka [43]

In this proof, the general case of the Craig-Sakamoto Theorem is shown by using cumulants, but first, let us recall some facts about them. A cumulant generating function is equal to the natural logarithm of the moment generating function of a variable y and we will denote the hth cumulant by $\kappa_h(y)$. Moreover $\kappa_h(sy) = s^h \kappa_h(y)$ for any constant s and the two random variables y_1 and y_2 are independent if and only if $\kappa_h(sy_1 + ty_2) = \kappa_h(sy_1) + \kappa_h(ty_2)$ for all integers h and all real numbers s and t. As shown at the end of Chapter 2, the jth cumulant of the quadratic form $x'Cx$, where $x \sim \mathrm{N}(\mu, V)$, is:

$$\kappa_j(C|\mu, V) = 2^{h-1}(h-1)!\mathrm{tr}(CV)^n h + n\mu' C(VC)^{h-1}\mu. \tag{3.89}$$

To prove the necessity part of the Craig-Sakamoto Theorem, Driscoll and his co-authors used a system of homogeneous linear equations $\Lambda\nu = 0$ involving the eigenvalues of sAV, tBV, and $(sA + tB)V$, with Λ the matrix of coefficients and ν, the vector of unknowns.

Let s and t be fixed, but arbitrary, and let $\{\lambda_1, \ldots, \lambda_k\}$ denote the union set of nonzero eigenvalues belonging to either sAV, tBV, or $(sA+tB)V$. Let C represent either sAV, tBV, or $(sA+tB)V$, then the multiplicity of λ_i being an eigenvalue of CV is denoted as $m_{i,C}$, the projection matrix as $P_{i,C}$ and we write $\mu_{i,C} = \mu' CP_{i,C}\mu$ and $(CV)^h = \sum_i \lambda_i^j P_{i,C}$. So the hth cumulant of the quadratic form $x'Cx$ becomes

$$\kappa_j(C|\mu, V) = 2^{h-1}(h-1)! \sum_{i=1}^{k} \lambda_i^h m_{i,C} + j\lambda_i^{h-1}\mu_{i,C}. \tag{3.90}$$

By the assumption of independence, $\kappa_h(sy_1 + ty_2) = \kappa_h(sy_1) + \kappa_h(ty_2)$, where y_1 and y_2 are random variables; then equation (3.90) becomes

$$\sum_{i=1}^{k} \lambda_i^h m_{i,sA+tB} - m_{i,sA} - m_{i,tB} + \sum_{i=1}^{k} h\lambda_i^{n-1}\mu_{i,sA+tB} - \mu_{i,sA} - \mu_{i,tB} = 0 \tag{3.91}$$

for all h. The first $2k$ of these equations have the matrix form

$$\Lambda\nu = 0, \tag{3.92}$$

where Λ is a $2k \times 2k$ matrix with elements $\Lambda_{i,j} = \lambda_j^i$ and $\Lambda_{i,k+j} = i\lambda_j^i$ for $i = 1, ..., 2k$ and $j = 1, ..., k$; and where ν is a $2k$-vector with elements $\nu_i = m_{i,sA+tB} - m_{i,sA} - m_{i,tB}$ and $\nu_{k+i} = \mu_{i,sA+tB} - \mu_{i,sA} - \mu_{i,tB}$ for $i = 1, ..., k$.

To prove that the vector $\nu = 0$, we will show that the matrix Λ is nonsingular and thus the only solution to this system of equations would be $\nu = 0$. First, we multiply columns $k+1, k+2, ..., 2k$ of Λ by $\lambda_1, \lambda_2, ..., \lambda_k$ respectively and permute rows and columns to obtain

$$L_k(\lambda_1, ..., \lambda_k) = \begin{pmatrix} \lambda_1^{2k} & 2k\lambda_1^{2k} & \cdots & \lambda_k^{2k} & 2k\lambda_k^{2k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_1^2 & 2k\lambda_1^2 & \cdots & \lambda_k^2 & 2k\lambda_k^2 \\ \lambda_1 & 2k\lambda_1 & \cdots & \lambda_k & 2k\lambda_k \end{pmatrix}. \tag{3.93}$$

The determinant is found by mathematical induction and proved to be equal to

$$|L_k| = (-1)^k \prod_i \lambda_i^3 \prod_{i<j} (\lambda_i - \lambda_j)^4 \qquad i = 1, ..., k. \tag{3.94}$$

By direct computation, we find

$$|L_1| = -\lambda_1^3; \qquad |L_2| = \lambda_1^3\lambda_2^3(\lambda_1 - \lambda_2)^4. \tag{3.95}$$

Assume that (3.94) holds for L_{k-1} with the above equalities for $|L_1|$ and $|L_2|$ establishing the induction base. Let $k \leq 3$ and $1 \geq m \geq k$ where m is fixed but arbitrary. By permutation of columns 1 and $2m-1$ and then columns 2 and $2m$, the matrix is then partitioned between the second and third rows and the second and third columns as

$$L_k(\lambda_m, \lambda_2, ..., \lambda_{m-1}, \lambda_1, \lambda_{m+1}, ..., \lambda_k) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3.96}$$

say, where D is the matrix $L_{k-1}(\lambda_2, ..., \lambda_{m-1}, \lambda_1, \lambda_{m+1}, ..., \lambda_k)$. By hypothesis, D is nonsingular and we have

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \cdot |A - CD^{-1}B| = |A - CD^{-1}B|(-1)^k \prod_{i \neq m} \lambda_i^3 \prod_{i<j, i\neq m, j\neq m} (\lambda_i - \lambda_j)^4. \tag{3.97}$$

Since m was fixed but arbitrary, it follows that $|L_k|$ has the factors $(-1)^{k-1}.\lambda_i^3$ for $1 \leq i \geq k$, and $(\lambda_i - \lambda_j)$ for $1 \leq i < j \leq k$. The power of the determinant of L_{k-1} is $4(k-1)$ by hypothesis, and if we add the powers of the above factors, we obtain $3 + 4(k-1) = 4k - 1$. In addition, we know that $|L_k|$ has degree $4(k-1)$ as a polynomial in any λ_i because that is the largest exponent attainable by multiplying elements in the different rows and columns. Hence, the only factor in $|L_k|$ missing must be a constant, let us call it c_k. If $m = 1$ we get

$$|A - CD^{-1}B| = c_k\lambda_1^3 \prod_{j=2}^{k} (\lambda_1 - \lambda_j)^4. \tag{3.98}$$

The coefficient of λ_1^{4k-1} on the right-hand side of (3.98) is -1. Hence $c_k = -1$ and we have proven (3.94). Because all λ_i are distinct and nonzero, the matrix L_k is nonsingular and we conclude by equation (3.90) that $m_{i,sA+tB} - m_{i,sA} - m_{i,tB} = 0$ and $\mu_{i,sA+tB} - \mu_{i,sA} - \mu_{i,tB} = 0$ for $i = 1, \ldots, k$. Therefore

$$\begin{aligned} m_{i,sA+tB} &= m_{i,sA} - m_{i,tB} & (3.99)\\ \mu_{i,sA+tB} &= \mu_{i,sA} - \mu_{i,tB} & (3.100) \end{aligned}$$

and we obtain the following equalities

$$\begin{aligned} \mathrm{tr}(sA+tB)V^h &= s^h \mathrm{tr}(AV)^h + t^h \mathrm{tr}(BV)^h & (3.101)\\ \mu'(sA+tB)V(sA+tB)^{h-1}\mu &= s^n \mu' A(VA)^{h-1}\mu + t^h \mu' B(VB)^{h-1}\mu & (3.102) \end{aligned}$$

for all h and, because they were arbitrary, for all s and t. Expanding the coefficient of s^2t^2 in (3.101), using $h = 4$, and writing $V = TT'$, we have[17]

$$\begin{aligned} \mathrm{tr}(AVBV + BVAV)^2 + 2\mathrm{tr}(AVBV)(AVBV)' &= 0\\ \mathrm{tr}(AVBTT' + BVATT')^2 + 2\mathrm{tr}(AVBTT')(AVBTT')' &= 0\\ \mathrm{tr}(T'AVBT + T'AVBT)^2 + 2\mathrm{tr}(T'AVBT)(T'AVBT)' &= 0. & (3.103) \end{aligned}$$

Both terms on the left-hand side of (3.103) are in general nonnegative, and since they are here required to add up to 0 they must each be equal to 0. Hence $T'AVBT = 0$ and $VAVBV = 0$. It follows from this and from (3.102) using $h = 4$ that

$$(T'AVB\mu)'(T'AVB\mu) + (T'BVA\mu)'(T'BVA\mu) = 0, \qquad (3.104)$$

and hence $VAVB\mu = 0$ and $VBVA\mu = 0$. Finally, setting $h = 2$ one has $\mu'AVB\mu = 0$. □

With this proof, Driscoll and Krasnicka (1995) updated their account of the development of the Craig-Sakamoto Theorem and added a section on the conditions found by Kawada (1950) [87][18] which are weaker than independence.

Then in 1992, Mathai and Provost [135] devoted an entire book to the distribution of quadratic forms in normal variables and provided valuable details concerning the independence of two quadratic forms. The proof presented there for the $\mathrm{N}(0, I)$ case is similar that by Ogasawara and Takahashi (1951); the proofs by Kawada (1950), Aitken (1950) and Lancaster (1954) are also mentioned. To show the case where the dispersion matrix $V \neq I$ but positive definite, they used the same transformation introduced by Aitken (1950). In the case where V is only nonnegative definite, they let $V = TT'$ and symmetrized the moment generating function obtaining the form

$$|I - 2sT'AT - 2tT'BT|^{-1/2} = |I - 2sT'AT|^{-1/2} \cdot |I - 2tT'BT|^{-1/2}$$

to adapt it to their proof for $\mathrm{N}(0, I)$. To show that the theorem holds for V positive definite and $x \sim \mathrm{N}(\mu, V)$, Mathai and Provost [135] presented a very short version of the proofs by Ogawa

[17] Cf. Ogasawara and Takahashi's proof for the $\mathrm{N}(0, I)$ case, subsection 3.2.5, and Kawada's Trace Lemma (our Lemma 18 above).

[18] Cf. Kawada's proof in subsection 3.2.2 of this Chapter.

(1950) and by Laha (1956), but without proving Laha's Lemma (our Lemma 18 above). As for the general case with V nonnegative definite, they used the very same lemma to demonstrate that the independence of quadratic forms is equivalent to the conditions obtained by Ogasarawa and Takahashi (1951). These two last proofs are short and again omit a proof of Laha's Lemma. See also the proofs by Provost (1994, 1996) which we will discuss below in Section 3.4.4.

Our next proof was published in 1993 by Ogawa [165][19].

3.4.3. Proof by Ogawa (1993)

We will show that if for real symmetric matrices A and B, cf. (2.1) in Chapter 2,

$$|I - sA - tB| = |I - sA| \cdot |I - tB| \qquad \forall \text{ real } s \text{ and } t, \tag{3.105}$$

then $AB = 0$.

If we put $s = t = \lambda$ in (3.105), then we obtain

$$|\lambda I_n - A| \cdot |\lambda I_n - (+B)| = \lambda^n |\lambda I_n - (A + B)|, \tag{3.106}$$

cf. (3.12) in Chapter 3, while if we put $s = -t = \lambda$ in (3.105), then we have

$$|\lambda I_n - A| \cdot |\lambda I_n - (-B)| = \lambda^n |\lambda I_n - (A - B)|. \tag{3.107}$$

From (3.106) we see that the set of nonzero eigenvalues of $A + B$ is the union of the sets of nonzero eigenvalues of A and nonzero eigenvalues of B, while from (3.107) we see that the set of nonzero eigenvalues of $A - B$ is the union of the sets of nonzero eigenvalues of A and of nonzero eigenvalues of $-B$. Hence

$$\operatorname{tr}((A + B)^4) = \operatorname{tr}(A^4) + \operatorname{tr}(B^4) = \operatorname{tr}((A - B)^4), \tag{3.108}$$

and so

$$\begin{aligned} 4\operatorname{tr}(A^2B^2) + 4\operatorname{tr}(A^3B) + 2\operatorname{tr}(ABAB) + 4\operatorname{tr}(AB^3) &= \\ 4\operatorname{tr}(A^2B^2) - 4\operatorname{tr}(A^3B) + 2\operatorname{tr}(ABAB) - 4\operatorname{tr}(AB^3) &= 0. \end{aligned} \tag{3.109}$$

Hence

$$4\operatorname{tr}(A^2B^2) + 2\operatorname{tr}(ABAB) = 0,$$

and $AB = 0$ follows at once from Kawada's Trace Lemma (Lemma 18 in our Chapter 3) and our proof is complete. □

Our next proof is by Provost [175, 176] and uses diagonalization and properties of determinants.

---

[19] I am very grateful to Professor Junjiro Ogawa for drawing this proof to my attention.

3.4.4. Proofs by Provost (1994, 1996)

To show that if for real symmetric matrices A and B, cf. (3.105) above,

$$|I - sA - tB| = |I - sA| \cdot |I - tB| \qquad \forall \text{ real } s \text{ and } t, \tag{3.110}$$

then $AB = 0$, Provost [175, 176] chose s so that $|\text{ch}(sA)| < 1$ and hence $I - sA$ is nonsingular. He then substituted $K = (I - sA)^{-1}$ in (3.110), which then becomes

$$|I - tKB| = |I - tB| \qquad \forall \text{ real } t. \tag{3.111}$$

Let G be an orthonormal matrix such that $G'AG = D = \text{diag}(\alpha_1, ..., \alpha_a, 0, ..., 0)$, where a is the rank of A and $\alpha_1, \ldots, \alpha_a$ are the nonzero eigenvalues of A. Then, using the geometric series expansion for $K = (I - sA)^{-1}$, cf. Lemma 7 in our Chapter 1, we obtain

$$|G'| \cdot |I - \sum_{k=0}^{\infty} (sA)^k tGG'B| \cdot |G| \equiv |G'| \cdot |I - tB| \cdot |G|. \tag{3.112}$$

We set

$$G'BG = H = \begin{pmatrix} H_{11} & H_{12} \\ H'_{12} & H_{22} \end{pmatrix}, \tag{3.113}$$

where H_{11} is a $a \times a$. We now observe that

$$G'A^kG = G'AGG'AGG' \cdots GG'AG = (G'AG)^k = \text{diag}(\alpha_1^k, ..., \alpha_a^k, 0, ..., 0). \qquad k = 1, \ldots \tag{3.114}$$

and so, using Lemma 6 in Chapter 1, we see that

$$\begin{aligned} G'(I - sA)^{-1}G &= G' \sum_{k=0}^{\infty} (sA)^k G = \sum_{k=0}^{\infty} s^k G'A^kG \\ &= I + \sum_{k=1}^{\infty} s^k \text{diag}(\alpha_1^k, ..., \alpha_a^k, 0, ..., 0) = \text{diag}(f^{(1)}, ..., f^a, 1, ..., 1), \end{aligned} \tag{3.115}$$

where $f^{(j)} = \sum_{k=0}^{\infty} (s\alpha_j)^k$, $j = 1, ..., a$. Writing $H = G'BG$, we obtain

$$|I - tH\text{diag}(f^{(1)}, ..., f^{(a)}, 1, ..., 1)| \equiv |I - tH|. \tag{3.116}$$

So by partitioning H is terms of H_{11}, H_{12}, H'_{12}, and H_{22}, we have

$$\begin{vmatrix} I_a - tH_{11}\text{diag}(f^{(1)}, ..., f^{(a)}) & -tH_{12} \\ tH'_{12}\text{diag}(f^{(1)}, ..., f^{(a)}) & I_{p-a} - tH_{22} \end{vmatrix} \equiv \begin{vmatrix} [I_a - tH_{11}] & -tH_{12} \\ tH'_{12} & [I_{p-a} - tH_{22}] \end{vmatrix}, \tag{3.117}$$

or equivalently

$$\begin{aligned} &|I - tH_{22}| \cdot |I - tH_{11}\text{diag}(f^{(1)}, ..., f^{(a)}) - tH_{12}(I - tH_{22})^{-1}tH'_{12}\text{diag}(f^{(1)}, ..., f^{(a)})| \\ &\qquad \equiv |I - tH_{22}| \cdot |(I - tH_{11}) - tH_{12}(I - tH_{22})^{-1}tH'_{12}|. \end{aligned} \tag{3.118}$$

Assuming that $|t|$ is small enough so that $I - tH_{22} > 0$, we may cancel $|I - tH_{22}|$ on both sides of (3.118). Postmultiplying by the determinant of $D_a = \mathrm{diag}(1 - t\alpha_1, ..., 1 - t\alpha_a)$, and writing

$$W = tH_{11} + tH_{12} \sum_{k=0}^{\infty} (tH_{22})^k tH_{12}', \tag{3.119}$$

we have

$$|D_a - W| \equiv |I - W| \prod_{j=1}^{a} (1 - t\alpha_j). \tag{3.120}$$

Comparing the coefficients of $f^{(a)}$, we obtain

$$\prod_{j=1}^{a}(-t\alpha_j) \quad \equiv \quad |I - W| \prod_{j=1}^{a}(-t\alpha_j) \quad \Rightarrow \quad -\log|I - W| = \sum_{h=1}^{\infty} \frac{\mathrm{tr}(W^h)}{h} = 0. \tag{3.121}$$

The coefficient of t^2 is

$$\mathrm{tr}(H_{11}^2) + \mathrm{tr}(H_{12}H_{12}'), \tag{3.122}$$

found by letting $h = 2$ in (3.121) and by letting $h = 1$ in (3.121) and $k = 0$ in (3.119). This coefficient must be 0 and so both H_{11} and H_{12} are zero[20]. Therefore

$$2G'BG = \begin{pmatrix} 0 & 0 \\ 0 & H_{22} \end{pmatrix} \tag{3.123}$$

and $G'BGG'AG = 0$, so $G'BAG = 0$ and $BA = 0$ which is equivalent to $AB = 0$. This completes Provost's proof. □

In 1994 Provost [175] also showed that the proofs by Searle [209] and Krafft [108] were not correct; he showed that $\mu' AVB\mu = 0$ holds for a specific μ and not for all μ. In addition, Provost [175] gave another proof based on some properties of the trace. In 1996 Provost [176] gave proofs for the simple case $\mathrm{N}(0, I)$, the general case $\mathrm{N}(\mu, V)$, with V positive definite, and for linear forms with $x \sim \mathrm{N}(0, I)$.

3.5. Overview: 1943–1996

In conclusion we believe that:

- The Craig-Sakamoto Theorem was first stated in 1943 by Craig [35] for $\mathrm{N}(0, I)$ and in 1944 by Sakamoto [196] for $\mathrm{N}(0, V)$, with V positive definite
- In 1944 Hotelling [73], attempting to complete the proof by Craig [35], made a "subtle gap" as explained by Driscoll and Gundberg (1986); see also Ogawa (1949, 1993) and Ogawa and Olkin (1997)

[20] As seen in many earlier proofs the trace of a nonnegative definite matrix is always nonnegative and zero if and only if the matrix itself is zero.

- The 1946 proof by Ogawa [158] for $N(0, I)$ is incomplete as pointed out by S. Nabeya; see also Ogawa (1949, 1993) and Ogawa and Olkin (1997)
- The 1949 proof by Matusita [138] is the first complete proof of the Craig-Sakamoto Theorem and is for $N(0, V)$ with V positive definite
- The 1949 proof by Ogawa [162] for $N(0, I)$ is complete
- In 1950 Kawada [87] gives a new complete proof of independence using results by Matérn (1949) on uncorrelatedness and presents a useful "Trace Lemma"
- The 1950 proof by Carpenter [27] for $N(\mu, I)$ refers to Craig (1943) and Hotelling (1944) for $N(0, I)$ and so the necessity part is incomplete; the proof of "Laha's Lemma" is adequate—see also Ogawa (1950, 1993) and Ogawa and Olkin (1997)
- The 1950 proof by Aitken [3] for $N(0, V)$, with V positive definite, is complete
- The 1950 proof by Ogawa [164] for $N(\mu, V)$, with V positive definite, is complete (and includes a complete proof of "Laha's Lemma" for real variables)
- In 1951 Ogasawara and Takahashi [157] gives first complete proof for the most general case $N(\mu, V)$, with V possibly singular; in addition, a relatively short proof is given for $N(0, I)$
- In 1956 Laha [110] introduced a different proof of the Craig-Sakamoto Theorem for $N(\mu, I)$, and extended the result to second-degree polynomials and to bilinear forms; presents "Laha's Lemma" (for complex variables), but without proof
- In 1958 Taussky [231] used pairs of matrices with property L to prove the necessity part of the Craig-Sakamoto Theorem for $N(0, I)$ and observed that the result still holds when the matrices are complex normal
- In a 1962 lecture Pao-Lu Hsu [74] presented a theorem on eigenvalues that is useful in proving the necessity part of the Craig-Sakamoto Theorem for $N(0, I)$
- The 1971 proof for $N(\mu, V)$, with V possibly singular, in Searle's *Linear Models* book is incomplete, but is completed in Searle's 1984 detailed class notes [208].
- The 1978 proof by Krafft [108] for $N(\mu, V)$, with V possibly singular, is incomplete
- The 1985 proof by Zielinski [248] for $N(\mu, V)$, with V possibly singular, is short but not completely clear
- The surveys by Reid and Driscoll (1988) and by Driscoll and Krasnicka (1995) include complete proofs for $N(\mu, V)$, with V possibly singular
- New proofs by Ogawa (1993) and by Provost (1994, 1996) for $N(\mu, V)$, with V possibly singular appear to be complete.

Chapter 4

Recent Developments: 1997–2000

In this last chapter, we introduce new proofs presented in the last few years. Some of these proofs are unpublished; to complete the current picture on the Craig-Sakamoto Theorem, however, we believe that they should appear in this thesis.

A very different proof to the ones already in existence was published in 1997 by Harville and Kempthorne [69] for second-degree polynomials (nonhomogeneous quadratic forms) in the most general case with nonzero mean vector μ and with dispersion matrix V possibly singular. This proof requires a lemma similar to Laha's Lemma but uses only polynomials in one variable instead of two as in Laha's Lemma and thus confirms the validity of the Craig-Sakamoto Theorem for quadratic forms and second-degree polynomials.

4.1. Proof by Harville and Kempthorne (1997)

Theorem 27 (Two second-degree polynomials with V possibly singular) *Let x be an $n \times 1$ random vector whose distribution is $\mathrm{N}(\mu, V)$, with V being of rank $r \leq n$, and let $q_1 = 2a'x + x'Ax$ and $q_2 = 2b'x + x'Bx$, where a and b are $n \times 1$ nonrandom vectors and A and B are $n \times n$ nonrandom symmetric matrices and $a^* = a + A\mu$ and $b^* = b + B\mu$. Then, q_1 and q_2 are independent if and only if*

$$
\begin{aligned}
VAVBV &= 0 && (4.1)\\
VAVb^* &= 0 && (4.2)\\
VBVa^* &= 0 && (4.3)\\
a^{*\prime}Vb^* &= 0 && (4.4)
\end{aligned}
$$

are satisfied.

Proof of Sufficiency. Suppose that the above equalities hold. Since V is nonnegative definite, we may write $V = L'L$, where L is $r \times n$, with $r = \text{rank}(V) \leq n$. It then follows easily (using e.g., Lemma 3 in Chapter 1 of Searle [209]), that (4.1)–(4.3) are equivalent to

$$LAVBL' = 0, \qquad LAVb^* = 0, \qquad \text{and} \qquad LBVa^* = 0. \tag{4.5}$$

Using Lemma 5 in Chapter 1, there exists an $r \times 1$ vector $z \sim \text{N}(0, I)$ such that $x = \mu + L'z$. Let us write $g = La^*$, $h = Lb^*$, $G = LAL'$, and $H = LBL'$. The $(r+1) \times 1$ vectors

$$z_1 = \begin{pmatrix} g'z \\ Gz \end{pmatrix} = \begin{pmatrix} g' \\ G \end{pmatrix} z \qquad \text{and} \qquad z_2 = \begin{pmatrix} h'z \\ Hz \end{pmatrix} = \begin{pmatrix} h' \\ H \end{pmatrix} z \tag{4.6}$$

are uncorrelated if and only if they are independent. The cross-covariance matrix of z_1 and z_2 is

$$\begin{pmatrix} g' \\ G \end{pmatrix} \begin{pmatrix} h' \\ H \end{pmatrix}' = \begin{pmatrix} g' \\ G \end{pmatrix} \begin{pmatrix} h & H \end{pmatrix} = \begin{pmatrix} g'h & g'H \\ Gh & GH \end{pmatrix} = \begin{pmatrix} a^{*\prime}Vb^* & a^{*\prime}VBL' \\ LAVb^* & LAVBL' \end{pmatrix} = 0$$

and this is equivalent to (4.5) and (4.4) and thus to (4.1)–(4.4). Let

$$q_1^* = 2g'z + z'Gz \qquad \text{and} \qquad q_2^* = 2h'z + z'Hz. \tag{4.7}$$

Since

$$q_1 = 2a'x + x'Ax = 2a'(\mu + L'z) + (\mu + L'z)'A(\mu + L'z) = 2a'\mu + \mu'A\mu + q_1^* \tag{4.8}$$

$$q_2 = 2b'x + x'Bx = 2b'(\mu + L'z) + (\mu + L'z)'B(\mu + L'z) = 2b'\mu + \mu'B\mu + q_2^*, \tag{4.9}$$

it follows that q_1 and q_2 are independent if and only if q_1^* and q_2^* are independent.

We may, however, express q_1^* and q_2^* as (homogeneous) quadratic forms in z_1 and z_2:

$$q_1^* = z_1' \begin{pmatrix} 0 & g'G^- \\ G^-g & G^- \end{pmatrix} z_1 \quad \text{and} \quad q_2^* = z_2' \begin{pmatrix} 0 & h'H^- \\ H^-h & H^- \end{pmatrix} z_2, \tag{4.10}$$

where G^- and H^- denote, respectively, symmetric generalized inverses of G and H, so that $GG^-G = G$ and $HH^-H = H$. Hence independence of z_1 and z_2 implies independence of q_1^* and q_2^*, and our proof of sufficiency is complete. □

Proof of Necessity. In this part of the proof, we need the following result:

Lemma 28 *Let $r_1(x), s_1(x)$ and $s_2(x)$ represent polynomials (with real coefficients) in a real variable x. Let*

$$r_2(x) \quad = \quad \gamma(x - \lambda_1)^{m_1} \dots (x - \lambda_k)^{m_k}, \tag{4.11}$$

where k is a nonnegative integer, $m_1, \dots, m_k$ are all integers strictly greater than zero, $\gamma \neq 0$, and $\lambda_1, \dots, \lambda_k$ are real numbers. Assume that

$$\log \frac{s_1(x)}{s_2(x)} = \frac{r_1(x)}{r_2(x)} \tag{4.12}$$

for all but the roots of $s_2(x)$ and $r_2(x)$. Then there exists a real number α such that $r_1(x) = \alpha r_2(x)$ and $s_1(x) = e^\alpha s_2(x)$ for all x.

A proof of this lemma is given by Harville and Kempthorne [69] but will be omitted here.

To prove necessity we note that q_1 and q_2 are independent if and only if q_1^* and q_2^* are independent, where q_1^* and q_2^* are as defined in (4.7) above. Let $m(\cdot,\cdot)$ be the joint moment generating function of q_1^* and q_2^*. Let c and d be positive scalars such that $I - 2tG - 2uH$ is positive definite for any t and u where $|t| < c$ and $|u| < d$, thus implying that $I - 2tG$ is positive definite for any t where $|t| < c$, and $I - 2uH$ is positive definite for any u where $|u| < d$. Hence, with $|t| < c$ and $|u| < d$, the moment generating function becomes:

$$\phi(t,u) = |I - 2tG - 2uH|^{-1/2}\exp\{2(tg+uh)'(I - 2tG - 2uH)^{-1}(tg+uh)\}. \tag{4.13}$$

Now, assuming that q_1 and q_2 are independent, then q_1^* and q_2^* are independent as well implying that $\phi(t,u) = \phi(t,0)\phi(0,u)$ and so,

$$\log\frac{|I - 2tG - 2uH|}{|I - 2tG||I - 2uH|} = 4\left\{\frac{(tg+uh)'(tg+uh)}{|I - 2tG - 2uH||} - \frac{t^2g'g}{(|I - 2tG|)} - \frac{u^2h'h}{(|I - 2uH|)}\right\}. \tag{4.14}$$

We know that

$$|I - 2tG| = \prod_{i=1}^{r}(-\lambda_i)(t - \lambda_i^{-1}), \tag{4.15}$$

where $\lambda_1,\ldots,\lambda_r$ are the nonzero eigenvalues of $2G$. We set $(I - 2uH)^{-1} = P'P$, P being nonsingular and $\tau_1,\ldots,\tau_r$ the nonzero eigenvalues of $2PGP'$; we have

$$|I - 2tG - 2uH| = |I - 2uH|\prod_{i=1}^{r}(-\tau_i)(t - \tau_i^{-1}). \tag{4.16}$$

With u fixed, the determinants $|I - 2tG|$, $|I - 2uH|$, and $|I - 2tG - 2uH|$ are polynomials in t only. Since every element of the inverse of a matrix can be expressed as the ratio of a cofactor and the determinant, we have

$$\log\frac{|I - 2tG - 2uH|}{|I - 2tG||I - 2uH|} = \frac{f_1(t,u)}{f_2(t,u)}, \tag{4.17}$$

where $f_1(t,u)$ is a polynomial in t and $f_2(t,u) = |I - 2tG||I - 2uH||I - 2tG - 2uH|$. Thus applying Lemma 28 with $x = t$, $s_1(t) = |I - 2tG - 2uH|$, $s_2(t) = |I - 2tG||I - 2uH|$, $r_1(t) = f_1(t,u)$, and $r_2(t) = f_2(t,u)$, yields $f_1(t,u) = \alpha f_2(t,u)$ and that

$$|I - 2tG - 2uH| = e^{\alpha}|I - 2tG||I - 2uH|. \tag{4.18}$$

Now, setting $\alpha = 0$ in (4.18), we have $e^{\alpha} = 1$ and hence

$$|I - 2tG - 2uH| = |I - 2tG||I - 2uH| \tag{4.19}$$

and $f(t,u) = 0$. So $GH = 0$, which implies that $VAVBV = L'GHL = 0$.

It can be shown that

$$\frac{\partial^2 f}{\partial t\partial u}_{|t=u=0} = 8g'h \tag{4.20}$$

and that

$$\frac{\partial^4 f}{\partial t^2\partial u^2}_{|t=u=0} = 64(Hg)'Hg + (Gh)'Gh. \tag{4.21}$$

Thus, we obtain $g'h = 0, Gh = 0, Hg = 0$, and conclude that

$$a^* V b^* = a^* L' L b^* = g' h = 0 \tag{4.22}$$

$$V A V b^* = L' L A L' L b^* = L' G h = 0 \tag{4.23}$$

$$V B V a^* = L' L B L' L a^* = L' H g = 0, \tag{4.24}$$

and our proof is complete. □

4.2. Olkin's 1997 Proof using Determinants

Also in 1997, a very different proof for the simplest case was provided by Olkin [169] using properties of determinants and the following result due to Sylvester, cf. e.g., Aitken [2, p. 87]. See also Marcus [132].

Lemma 29 (Sylvester's Result) *Let $C(i_1, ..., i_k)$ denote the determinant of a principal submatrix of $C_{n\times n}$ with rows and columns $i_1, ..., i_k$ and let $D = \operatorname{diag}(x_1, ..., x_n)$ be a diagonal matrix. Then*

$$|D + C| = |C| + \sum x_1 C(2, ..., n) + \sum x_1 x_2 C(3, ..., n) + \cdots + \sum x_1 \cdots x_{n-1} C(n) + \prod x_i \tag{4.25}$$

holds.

Proof. Assuming independence and, without loss in generality, that $A = D_a = \operatorname{diag}(a_1, ..., a_n)$ is diagonal, and replacing all the negative signs by positive signs in $|I - sA - tB| = |I - sA||I - tB|$, we get

$$|I + sA + tB| = |I + sA||I + tB|. \tag{4.26}$$

Let $x_i \equiv 1 + sa_i$ and $C \equiv tB$; then using Lemma 29, we obtain

$$|D_a + C| = |C| + \sum x_1 C(2, ..., n) + \sum x_1 x_2 C(3, ..., n) + \cdots + \sum x_1 \cdots x_{n-1} C(n) + \prod x_i, \tag{4.27}$$

where $\sum x_1, ..., x_k C(k, ..., n) = \sum_{i<j} x_1, ..., x_k C(k, ..., n)$. Putting (4.27) into (4.26) gives

$$\begin{aligned} t^n |B| + t^{n-1} \sum x_1 B(2, ..., n) + \cdots + t \sum x_1, ..., x_{n-1} b_{nn} \\ = \prod_1^n x_i \left(t^n |B| + t^{n-1} \sum B(2, ..., n) + ... + t \sum b_{nn} \right), \end{aligned} \tag{4.28}$$

which holds for all s and t. Rearranging (4.28) yields

$$t^n |B| \left(\prod_1^n x_i - 1 \right) + t^{n-1} \sum B(2, ..., n) \left(\prod_1^n x_i - x_1 \right) + \cdots + t \sum b_{nn} \left(\prod_1^n x_i - \prod_1^{n-1} x_i \right) = 0. \tag{4.29}$$

Examining (4.29), we find that

$$\sum b_{nn}\left(\prod_1^n x_i - \prod_1^{n-1} x_i\right) = 0 \tag{4.30}$$

and

$$\sum B(n-1,n)\left(\prod_1^n x_i - \prod_1^{n-2} x_i\right) = 0. \tag{4.31}$$

We know that

$$\prod_1^m x_i = 1 + s\sum a_1 + s^2\sum a_1 a_2 + \cdots + s^m \prod_1^m a_i; \tag{4.32}$$

and that for each $r = 1, 2, \ldots n-1$, $\prod_1^n x_i - \prod_1^r x_i$ is a polynomial in s of the form $sd_1 + s^2 d_2 + \cdots + s^n d_n$, where the coefficients are functions of $a_1, \ldots, a_n$ and depend on r. For example,

$$\prod_1^n x_i - x_n = (1 + sa_n)\left(s\sum_1^{n-1} a_i + s^2 \sum_1^{n-1} a_i a_j + \cdots + \prod_1^{n-1} a_i\right). \tag{4.33}$$

So we obtain

$$\prod_1^n x_i - \prod_1^{n-1} x_i = 0 \tag{4.34}$$

$$\prod_1^n x_i - \prod_1^{n-2} x_i = 0. \tag{4.35}$$

We notice that the left-hand sides of (4.34) and (4.35) are polynomials in s and vanish for all s. Hence each sum of products of the a_i on the right-hand side of (4.33) must be zero. In the case where each $a_1 = \cdots = a_n = 0$, we have $D_a B = 0$ and therefore $AB = 0$. If this is not the case, then there must exist a set of a_i such that

$$a_1 \neq 0, \ldots, a_r \neq 0, \quad a_{r+1} = \cdots = a_n = 0, \qquad 1 \leq r < n. \tag{4.36}$$

Hence $x_{r+1} = \cdots = x_n = 1$ and we let $\mathcal{A} = \{1, \ldots, r\}$ and $\mathcal{B} = \{r+1, \ldots, N\}$. For any subset of $\mathcal{B}$ we get

$$\prod_1^r x_i \prod_{j \in \mathcal{B}} x_j = \prod_1^r x_i. \tag{4.37}$$

The coefficients of s^m in (4.30) vanish for $m \in \mathcal{B}$. For $m = r$ all a_i vanish if they are in $\mathcal{B}$, and the coefficients of $b_{11}, \ldots, b_{rr}$ are respectively $a_1, \ldots, a_r$, which yields

$$(b_{11} + \cdots + b_{rr})\prod_1^r a_i = 0; \tag{4.38}$$

hence $b_{11} + \cdots + b_{rr} = 0$. As for (4.31), the term

$$\left(\sum_{r+1}^n B(1,j) + \cdots + \sum_{r+1}^n B(r,j) + \sum\sum_{i<j}^r B(i,j)\right) \times \sum_{i<j} a_i a_j \tag{4.39}$$

does not vanish giving

$$\sum_{j\in\mathcal{B}}\sum_{i\in\mathcal{A}} B(i,j) + \sum_{j\in\mathcal{A}}\sum_{i\in\mathcal{B}} B(i,j) = 0. \tag{4.40}$$

With the following lemma, our proof is complete.

Lemma 30 (Determinantal Result) *If* $b_{11} + \cdots + b_{rr} = 0$ *and*

$$\sum_{j\in\mathcal{B}}\sum_{i\in\mathcal{A}} B(i,j) + \sum_{j\in\mathcal{A}}\sum_{i\in\mathcal{A}} B(i,j) = 0 \tag{4.41}$$

then

$$B = \begin{pmatrix} 0 & 0 \\ 0 & B_{22} \end{pmatrix} \tag{4.42}$$

and, therefore, $D_a B = AB = 0$.

4.3. Proof by Drury, Dumais and Styan (1999)

This proof by Drury, Dumais and Styan (1999) is unpublished[1]. We believe that this proof is new; it is shorter than similar proofs given by Ogasawara and Takahashi (1951), Scarowsky (1973), Scarowsky and Styan (1982), and Ogawa (1993).

Let $K = (I - sA)^{-1}$ for those s so that $I - sA$ is invertible and $\lambda = 1/t$, $t \neq 0$. Then we may write

$$|I - sA - tB| = |I - sA| \cdot |I - tB| \tag{4.43}$$

as

$$|\lambda I - KB| = |\lambda I - B| \tag{4.44}$$

for all real λ. The characteristic polynomials of KB and B must, therefore, coincide and so their eigenvalues are equal. Hence, using the power series expansion of a matrix geometric series, (cf. our Lemma 6 in Chapter 1),

$$K = (I - sA)^{-1} = \sum_{j=0}^{\infty}(sA)^j \tag{4.45}$$

[1]Presented by Dumais at The Eighth International Workshop on Matrices and Statistics, Tampere, Finland, 6–7 August 1999, and by Styan in the "Special Session on the Interaction Between Statistics and Matrix Theory" at the Annual Meeting of the Statistical Society of Canada, Regina, Saskatchewan, 7–8 June 1999 and at the Conference on Functional Analysis and Linear Algebra, Indian Statistical Institute–Delhi Centre, 3–7 January 2000.

it follows that

$$\mathrm{tr}\{B^2\} = \mathrm{tr}\{(KB)^2\} = \mathrm{tr}\{(I-sA)^{-1}B(I-sA)^{-1}B\} = \mathrm{tr}\left\{\sum_{h=0}^{\infty}(sA)^h B \sum_{k=0}^{\infty}(sA)^k B\right\} \tag{4.46}$$

for all real s in an interval around 0. Since $\mathrm{tr}B^2$ does not involve s, it follows that the coefficient of s^2 on the right-hand side of (4.46) must be zero. Putting $(h,k) = (2,0), (1,1)$ and $(0,2)$ yields

$$\mathrm{tr}A^2B^2 + \mathrm{tr}ABAB + \mathrm{tr}BA^2B = 2\mathrm{tr}A^2B^2 + \mathrm{tr}(AB)^2 = 0. \tag{4.47}$$

From Kawada's Trace Lemma (our Lemma 18 in Chapter 3), it follows at once that $AB = 0$, and our proof is complete. □

4.4. Proof by Li (2000)

This new proof by Chi-Kwong Li in the paper [117][2] depends only on the following well-known fact:

Lemma 31 *Suppose $C = (c_{ij})$ is an $n \times n$ real symmetric matrix with the largest eigenvalue equal to λ_1. Then $c_{ii} \le \lambda_1$ for all $i = 1, \ldots, n$. If $c_{ii} = \lambda_1$, then $c_{ij} = 0 = c_{ji}$ for all $j \neq i$.*

For the sake of completeness, we give a short proof.

Proof. Suppose C satisfies the hypothesis of the lemma and the largest eigenvalue of C has multiplicity m with $1 \le m \le n$. Then there is an orthonormal basis $\{v_1, \ldots, v_n\}$ for $\mathbb{R}^n$ such that $Cv_j = \lambda_j v_j$ with $\lambda_1 = \cdots = \lambda_m > \lambda_{m+1} \ge \cdots \ge \lambda_n$. Let $\{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$. For any i with $1 \le i \le n$, there exist $t_1, \ldots, t_n \in \mathbb{R}$ with $\sum_{j=1}^n t_j^2 = 1$ such that $e_i = \sum_{j=1}^n t_j v_j$ and $c_{ii} = e_i^t C e_i = \sum_{j=1}^n t_j^2 \lambda_j \le \lambda_1$. The equality holds if and only if $t_{m+1} = \cdots = t_n = 0$, i.e., e_i is an eigenvector of C corresponding to the largest eigenvalue. Thus, $Ce_i = \lambda_1 e_i$, and hence $c_{ii} = \lambda_1$ is the only nonzero entry in the ith column. Since C is symmetric, c_{ii} is also the only nonzero entry in the ith row. □

We are now ready to present our proof of

Theorem 32 (The Craig-Sakamoto Theorem) *Two $n \times n$ real symmetric matrices A and B satisfy $AB = 0$ if and only if*

$$|I - sA - tB| = |I - sA| \cdot |I - tB| \quad \forall \textit{ real } s \textit{ and } t. \tag{4.48}$$

Proof. The ($\Leftarrow$) part is clear. We prove the converse by induction on n. The result is clear if $n = 1$. Suppose $n > 1$ and the result is true for symmetric matrices of sizes smaller than n. Let A and B be nonzero $n \times n$ real symmetric matrices satisfying (4.48). Denote by $\rho(C)$ the spectral radius of a square matrix C. Replacing A by $\pm A/\rho(A)$ and B by $B/\rho(B)$, we may

[2] To be published in 2000. Reprinted here with the kind permission of Chi-Kwong Li.

assume that $1 = \rho(A) = \rho(B)$ is the largest eigenvalue of A. Let Q be an orthogonal matrix such that $QAQ^t = I_m \oplus \mathrm{diag}(a_{m+1}, \ldots, a_n)$ with $1 > a_{m+1} \geq \cdots \geq a_n$. We shall show that $(QAQ')(QBQ') = 0$ and hence $AB = 0$.

For simplicity, we assume that $Q = I$. Let $t = \pm 1$. If $r > 1$, then both A/r and tB/r have eigenvalues in the open interval $(-1, 1)$. Thus, $I - A/r$ and $I - tB/r$ are invertible, and

$$|I - A/r - tB/r| = |I - A/r| \cdot |I - tB/r| \neq 0.$$

Moreover, since

$$|I - A - tB| = |I - A| \cdot |I - tB| = 0,$$

we see that 1 is the largest eigenvalue of the matrix $A + tB$ for $t = \pm 1$.

Next, we show that B is of the form $0_m \oplus B_2$. Note that all the first m diagonal entries of A are equal to the largest eigenvalue of $A \pm B$. If the first m diagonal entries of B are not all 0, then the matrix $A + B$ or $A - B$ will have a diagonal entry larger than 1, contradicting Lemma 31. So, all the first m diagonal entries of the matrix $A + B$ equal the largest eigenvalue. By Lemma 31 again, $A + B$ must be of the form $I_m \oplus C_2$. Hence, B is of the form $0_m \oplus B_2$, as asserted.

Now, let $A = I_m \oplus A_2$. Then for any real numbers s and t with $s \neq 1$, we have

$$|I_{n-m} - sA_2 - tB_2| = \frac{|I_n - sA - tB|}{|I_m - sI_m|} = \frac{|I_n - sA| \cdot |I_n - tB|}{|I_m - sI_m|} = |I_{n-m} - sA_2| \cdot |I_{n-m} - tB_2|.$$

By continuity, we can remove the restriction that $s \neq 1$. Using the induction assumption, we see that $A_2 B_2 = 0$. Hence, we have $AB = 0$ as desired. □

4.5. Extension by Li and Styan (2000) for Normal Matrices

Our last proof of the Craig-Sakamoto Theorem, by Li and Styan [118][3], assumes that A and B are complex normal matrices. Taussky [231] pointed out already in 1958 that the Craig-Sakamoto Theorem could be extended to complex normal matrices; her proof, however, relied on the so-called property L of a pair of (normal) matrices. Here, we show that such an extension can be done without using property L. We begin with the following extension to complex normal matrices of Hsu's Theorem (our Theorem 26 in Chapter 3).

Lemma 33 *Let A and B be $n \times n$ complex normal matrices. Suppose A, B and $A+B$ have nonzero eigenvalues (counting multiplicities) $\alpha_1, \ldots, \alpha_a$, $\beta_1, \ldots, \beta_b$, and $\alpha_1, \ldots, \alpha_a, \beta_1, \ldots, \beta_b$, respectively. Then $AB = 0$.*

Proof. Let $D_1 = \mathrm{diag}(\alpha_1, \ldots, \alpha_a)$ and $D_2 = \mathrm{diag}(\beta_1, \ldots, \beta_b)$. We may assume that $A = D_1 \oplus 0_{n-a}$. Otherwise, replace A and B by U^*AU and U^*BU for a suitable unitary U. Let

$$V = \begin{pmatrix} V_1 & V_2 & V_3 \\ V_4 & V_5 & V_6 \\ V_7 & V_8 & V_9 \end{pmatrix}$$

[3] Unpublished. Reproduced here with permission of Chi-Kwong Li and George P. H. Styan.

be unitary so that V_1 is $a \times a$, V_5 is $b \times b$ and

$$VBV^* = 0_a \oplus D_2 \oplus 0_{n-a-b}.$$

Consider

$$S = \begin{pmatrix} I_a & 0 & 0 \\ V_4 & V_5 & V_6 \\ 0 & 0 & 0_{n-a-b} \end{pmatrix}.$$

Then $A + B = S^* DS$ with $D = D_1 \oplus D_2 \oplus 0_{n-a-b}$. Note that the eigenvalues of XY and YX are the same for two square matrices X and Y. Thus the eigenvalues of $A + B = S^* DS$ are the same as those of

$$DSS^* = D \begin{pmatrix} I_a & V_4^* & 0 \\ V_4 & I_b & 0 \\ 0 & 0 & 0_{n-a-b} \end{pmatrix}.$$

Let $m = a + b$. Then the sum of all the $m \times m$ principal submatrices of $A + B$ is equal to the mth elementary symmetric function of the eigenvalues of $A + B$, which is $\prod_{j=1}^{a} \alpha_j \prod_{k=1}^{b} \beta_k$. Using the same arguments for DSS^* and the fact that the only nonzero $m \times m$ principal submatrix of DSS^* is the leading one, we conclude that

$$\prod_{j=1}^{a} \alpha_j \prod_{k=1}^{b} \beta_k = \prod_{j=1}^{a} \alpha_j \prod_{k=1}^{b} \beta_k \det \begin{pmatrix} I_a & V_4^* \\ V_4 & I_b \end{pmatrix} = \prod_{j=1}^{a} \alpha_j \prod_{k=1}^{b} \beta_k |I_r - V_4^* V_4|.$$

Hence $V_4^* V_4 = 0_a$ and thus

$$B = S^*(0_a \oplus D_2 \oplus 0_{n-a-b})S = 0_a \oplus \begin{pmatrix} V_5^* D_2 V_5 & V_5^* D_2 V_6 \\ V_6^* D_2 V_5 & V_6^* D_2 V_6 \end{pmatrix}.$$

It follows that $AB = 0$. $\square$

Note that if $m = n$, our proof is basically the same as that of Hsu's Theorem (our Theorem 26 in Chapter 3).

We are now ready to state and prove our main theorem.

Theorem 34 *Let the complex $n \times n$ normal matrices A and B have nonzero eigenvalues (counting multiplicities) $\alpha_1, \ldots, \alpha_a$, and $\beta_1, \ldots, \beta_b$, respectively. The following conditions are equivalent.*

(a) $AB = 0$.

(b) *There is a unitary matrix U such that $U^* AU = \text{diag}(\alpha_1, \ldots, \alpha_a, 0, \ldots, 0)$ and*

$$U^* BU = 0_a \oplus \text{diag}(\beta_1, \ldots, \beta_b, 0, \ldots, 0).$$

(c) *There are infinite sets $S, T \subseteq \mathbb{C}$ such that for any $(s,t) \in S \times T$, it follows that*

$$|I - sA - tB| = |I - sA| \cdot |I - tB|.$$

(d) *There exist nonzero $s, t \in \mathbb{C}$ such that $sA + tB$ has eigenvalues*

$$s\alpha_1, \ldots, s\alpha_a, t\beta_1, \ldots, t\beta_b, 0, \ldots, 0.$$

Proof. (a) $\Rightarrow$ (b): Suppose that $AB = 0$. Let U be unitary so that $U^*AU = \mathrm{diag}(\alpha_1, \ldots, \alpha_a) \oplus 0_{n-a}$. Suppose

$$U^*BU = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}.$$

Since $AB = 0$, we see that B_1 and B_2 are zero blocks. Since $BB^* = B^*B$, we see that B_3 is also a zero block. Suppose V is unitary such that $V^*B_4V = \mathrm{diag}(\beta_1, \ldots, \beta_b) \oplus 0_{n-a-b}$. Replace U by $U(I_a \oplus V)$. Then U^*AU and U^*BU are of the forms specified in (b).

(b) $\Rightarrow$ (c): Immediate.

(c) $\Rightarrow$ (d): Note that

$$\begin{aligned} |\lambda I - sA - tB| &= \lambda^n |I - (s/\lambda)A - (t/\lambda)B| \\ &= \lambda^n |I - (s/\lambda)A| \cdot |I - (t/\lambda)B| \\ &= \lambda^{n-a-b} \prod_{j=1}^{a} (\lambda - s\alpha_j) \prod_{k=1}^{b} (\lambda - t\beta_k) \end{aligned}$$

for all $(s/\lambda, t/\lambda) \in R \times S$. Thus, the polynomial

$$|\lambda I - sA - tB| \equiv \lambda^{n-a-b} \prod_{j=1}^{a} (\lambda - s\alpha_j) \prod_{k=1}^{b} (\lambda - t\beta_k).$$

and condition (d) follows.

(d) $\Rightarrow$ (a): Apply Lemma 33 to the matrices sA and tB to conclude that $AB = 0$, and our proof is complete. □

Several remarks are in order. The proof of (c) $\Rightarrow$ (d) actually reveals that A and B satisfy property L. One may then conclude that $AB = BA$ and prove that (d) $\Rightarrow$ (a) as in Taussky [231]. If A and B are real, then one may assume that U is orthogonal in condition (b), and that $s, t \in \mathbb{R}$ in conditions (c) and (d).

Bibliography

1. 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 [MR 1:346a, Zbl 027.23302]. [Corrigendum: *Proceedings of the Royal Society of Edinburgh Series A*, 62, 277 (1948).]
2. Aitken, A. C. (1944). On the number of distinct terms in the expansion of symmetric and skew determinants. *Edinburgh Mathematics*, Note 34, 1–5 [Zbl 060.03106, MR 6.32n].
3. Aitken, A. C. (1950). On the statistical independence of quadratic forms in normal variates. *Biometrika*, 37, 93–96 [MR 12:35j, Zbl 037.08902].
4. Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. Second Edition. Wiley, New York [MR 86b:62079].
5. Anderson, T. W.; Das Gupta, Somesh, and Styan, George P. H. (1977). *A Bibliography of Multivariate Statistical Analysis*. Reprint Edition. Krieger, Huntington, NY [MR 56:1585; Zbl 263.62001, 421.62033].
6. Anderson, T. W., and Fang, Kai-Tai (1987). Cochran's theorem for elliptically contoured distributions. *Sankhyā Series A*, 49, 305–315 [MR 91b:62094, Zbl 655.62057]. (Reprinted in [49], pp. 137–146.)
7. Anderson, T. W., and Styan, George P. H. (1982). Cochran's theorem, rank additivity and tripotent matrices. In *Statistics and Probability: Essays in Honor of C. R. Rao* (G. Kallianpur, P. R. Krishnaiah and J. K. Ghosh, eds.), North-Holland, Amsterdam, pp. 1–23 [MR 83h:15002, Zbl:62030]. (Reprinted in [226], pp. 1307–1329.)
8. Bakirov, N. K. (1989). Criteria for quadratic forms to be independent in two random variables [in Russian]. In *Studies in Approximation Theory: Collection of Articles* (R. S. Yulmukhametov, ed.). Institut Matematiki BNTs AN SSSR, Ufa, pp. 138–150 [Zbl 825.62362].
9. Baksalary, Jerzy K., and Hauke, Jan (1984). Inheriting independence and chi-squaredness under certain minus orderings. *Statistics & Probability Letters*, 2, 35–38 [Zbl 584.62075].
10. Baksalary, Jerzy K., and Hauke, Jan (1990). A further algebraic version of Cochran's theorem and matrix partial orderings. *Linear Algebra and its Applications*, 127, 157–169 [MR 91k:15043, Zbl 696.15005].
11. Baksalary, Jerzy K.; Hauke, Jan, and Kala, Radoslaw (1980). Nonnegative definite solutions to some matrix equations occuring in the distribution theory of quadratic forms. *Sankhyā Series A*, 42, 283–291 [MR 83e:15015, Zbl 501.62036].
12. Baksalary, J. K.; Hauke, J., and Styan, G. P. H. (1994). On some distributional properties of quadratic forms in normal variables and on some associated matrix partial orderings. In *Multivariate Analysis and Its Applications* (T. W. Anderson, K. T. Fang and I. Olkin, eds.), IMS Lecture Notes–Monograph Series, vol. 24, pp. 111–121 [MR 98e:62010].
13. Baldessari, Bruno (1967). The distribution of a quadratic form of normal random variables. *The Annals of Mathematical Statistics*, 38, 1700–1704 [MR 36:2241, Zbl 155.27301].
14. Baldessari, Bruno, and Gallo, Francesca (1980). On the joint distribution of the sample mean and variance of dependent normal data and related estimation problems. *Metron*, 38, 93–106 [Zbl 497.62044].

15. Banerjee, K. S. (1964). A note on idempotent matrices. *The Annals of Mathematical Statistics*, 35, 880–882 [MR 28:8071].

16. Banerjee, K. S., and Nagase, G. (1976). A note on the generalization of Cochran's theorem. *Communications in Statistics–A, Theory and Methods*, 9, 837–842 [MR 55:4513, Zbl 337.62034].

17. Bapat, R. B. (1999). *Linear Algebra and Linear Models.* Second Edition. Springer-Verlag, New York, and Hindustan Book Agency, New Delhi. (Original version: Hindustan Book Agency, New Delhi, 1993. [MR 95e:62071, Zbl 834.62062].)

18. Basu, J. P.; Odell, P. L., and Lewis, T. O. (1973). Distribution of certain quadratic forms in multivariate normal variates. *Industrial Mathematics*, 23, 27–35 [MR 48:9970].

19. Bhapkar, V. P. (1970). The invariance and distribution of the quadratic form $x'\Sigma x$. Technical Report No. 4, Dept. of Statistics, University of Kentucky, 8 pp.

20. 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 [Zbl 091.31702].

21. 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 [MR 29:6570, Zbl 121.14404].

22. Bhimasankaram, P., and Majumdar, Dibyen (1980). Hermitian and nonnegative definite solutions of some matrix equations connected with distribution of quadratic forms. *Sankhyā Series A*, 42, 272–282 [MR 83e:15014, Zbl 508.62042].

23. Binotti, Annetta Maria (1980). Sulla verifica dell'indipendenza stocastica tra forme lineari e forme quadratiche nell'ambito della tecnica BLUS [in Italian]. *Statistica (Bologna)*, 40, 329–335 [MR 82f:62118, Zbl 452.62053].

24. Brand, Louis (1969). On the product of singular symmetric matrices. *Proceedings of the American Mathematical Society*, 22, 377.

25. Cacoullos, T., and Koutras, M. (1984). Quadratic forms in spherical random variates: generalized noncentral χ^2 distribution. *Naval Research Logistics Quarterly*, 31, 447–461 [MR 86b:62026, Zbl 561.62049].

26. Cao, Zhen Hua (1998). The variance of quadratic forms with block matrices [in Chinese]. *Gongcheng Shuxue Xuebao/Journal of Engineering Mathematics (Xi'an)*, 15 (1), 51–55.

27. Carpenter, Osmer (1950). Note on the extension of Craig's theorem to non-central variates. *The Annals of Mathematical Statistics*, 21, 455–457 [MR 12:621e, Zbl 038.09604].

28. Chaganty, N. Rao, and Vaish, A. K. (1997). An invariance property of common statistical tests. *Linear Algebra and its Applications*, 264, 421–437 [MR 98m:62161, Zbl 980.01213].

29. Chipman, J. S., and Rao, M. M. (1964). Projections, generalized inverses, and quadratic forms. *Journal of Mathematical Analysis and Applications*, 9, 1–11 [Zbl 142.00401].

30. 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 [Zbl 009.12004].

31. Conradie, W. J., and Gupta, A. K. (1987). Quadratic forms in complex normal variates: basic results. *Statistica (Bologna)*, 47, 73–84 [MR 88i:62093, Zbl 619.62045].

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

33. Craig, A. T. (1936). A certain mean-value problem in statistics. *Bulletin of the American Mathematical Society*, 42, 670–674 [Zbl 015.11805].

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

35. Craig, Allen T. (1943). Note on the independence of certain quadratic forms. *The Annals of Mathematical Statistics*, 14, 195–197 [MR 5:127a, Zbl 060.3080].

36. Craig, Allen T. (1947). Bilinear forms in normally correlated variables. *The Annals of Mathematical Statistics*, 18, 565–573 [MR 9:294e].

37. Crofts, A. E., Jr. (1971). Some covariance relations for linear and quadratic forms. *Trabajos Estadística*, 22, 181–183 [MR 47:6008].

38. Das Gupta, S.; Ghosh, J. K.; Mitra, S. K.; Mukhopadhyay, A. C.; Rao, P. S. S. N. V. P., and Sarma, Y. R., eds. (1994). *Selected Papers of C. R. Rao*, Volume 2. Wiley Eastern, New Delhi [MR 95a:01029].

39. Dieulefait, E. (1951). Sobre las formas cuadraticas a variables aleatorias [in Spanish]. *Anales de la Sociedad Cientifica Argentina*, 151, 167–172 [Zbl 042.13602].

40. Dik, J. J., and de Gunst, M. C. M. (1985). The distribution of general quadratic forms in normal variables. *Statistica Neerlandica*, 39, 14–26 [MR 87d:62101, Zbl 591.62043].

41. Driscoll, Michael F. (1999). An improved result relating quadratic forms and chi-square distributions. *The American Statistician*, 53, 273–275.

42. Driscoll, Michael F., and Gundberg, William R., Jr. (1986). A history of the development of Craig's theorem. *The American Statistician*, 40, 65–70 [Zbl 589.62034].

43. Driscoll, Michael F., and Krasnicka, Barbara (1995). An accessible proof of Craig's theorem in the general case. *The American Statistician*, 49, 59–62 [MR 96d:62018].

44. Drury, S. W.; Dumais, Mylène, and Styan, George P. H. (1999). A new proof of the Craig-Sakamoto Theorem. Presented in a talk (by Dumais) at Eighth International Workshop on Matrix Methods for Statistics (Tampere, Finland, August 1999) and (by Styan) in the "Special Session on the Interaction Between Statistics and Matrix Theory" at the Annual Meeting of the Statistical Society of Canada, Regina, Saskatchewan, 7–8 June 1999 and at the Conference on Functional Analysis and Linear Algebra, Indian Statistical Institute–Delhi Centre, 3–7 January 2000.

45. Dumais, Mylène, and Styan, George P. H. (1998). A bibliography on the distribution of quadratic forms in normal variables, with special emphasis on the Craig-Sakamoto theorem and on Cochran's theorem. In *Three Bibliographies and a Guide* (George P. H. Styan, ed.), Seventh International Workshop on Matrices and Statistics, in celebration of T. W. Andersons 80th birthday (Fort Lauderdale, Florida, December 1998), pp. 1–9.

46. Eaton, Morris L. (1983). *Multivariate Statistics: A Vector Space Approach.* Wiley, New York [MR 86i:62086].

47. Fan, Jianqing (1986). Noncentral Cochran's theorem for elliptically contoured distributions. *Acta Mathematica Sinica New Series*, 2, 185–198 [MR 88d:62089]. (Reprinted in [49], pp. 177–191.)

48. Fang, Kai-Tai (1987). An introduction to generalized multivariate analysis: on the theory of elliptically contoured distributions [in Chinese, with Engish summary]. *Advances in Mathematics (Beijing)*, 16, 1–15 [MR 88c:62081, Zbl 615.62056].

49. Fang, Kai-Tai, and Anderson, T. W., eds. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions.* Allerton Press, New York [MR 91e:62006].

50. Fang, Kai-Tai; Fan, Jianqing, and Xu, Jian Lun (1987). The distributions of quadratic forms of random idempotent matrices and their applications [in Chinese]. *Yingyong Gailü Tongji/Chinese Journal of Applied Probability and Statistics (Shanghai)*, 3, 289–297 [MR 89g:62076, Zbl 689.62038].

51. Fang, Kai-Tai, and Wu, Yue Hua (1990). Distribution of quadratic forms and Cochran's theorem. In *Statistical Inference in Elliptically Contoured and Related Distributions* (Kai-Tai Fang and T. W. Anderson, eds.), Allerton Press, New York, pp. 147–161 [MR 91g:62040].

52. Fang, Kai-Tai, and Zhang, Yao-Ting (1990). *Generalized Multivariate Analysis* [in English]. Science Press, Beijing & Springer-Verlag, Berlin [MR 92f:62066, Zbl 724.62054]. (Expanded translation from the Chinese of Zhang & Fang (1983) [247]. Cf. Lemma 2.8.3, pp. 77–79.)

53. Fisher, Lloyd (1973). An alternative approach to Cochran's theorem. *The American Statistician*, 27, 109 [MR 50:1360].

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

55. Getson, Albert J., and Hsuan, Francis C. (1988). *{2}-Inverses and Their Applications in Statistics.* Lecture Notes in Statistics, vol. 47. Springer-Verlag, New York [MR 90e:15004, Zbl 671.62003].

56. 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 [MR 35:1106, Zbl 147.17402]. (Corrigenda: 28 (1966), 584 [MR 29:5341, Zbl 117.14103].)

57. Good, I. J. (1969). Conditions for a quadratic form to have a chi-squared distribution. *Biometrika*. 56, 215–216 [Zbl 179.24203]. [Correction: *Biometrika*, 57, 225 (1970).]

58. Graybill, Franklin A. (1961). *An Introduction to Linear Statistical Models: Volume I.* McGraw-Hill, New York [MR 23:A3612, Zbl 121.35605].

59. Graybill, Franklin A. (1976). *Theory and Application of the Linear Model.* Duxbury and Brooks/Cole [MR 56:13457].

60. Graybill, Franklin A., and Marsaglia, George (1957). Idempotent matrices and quadratic forms in the general linear hypothesis. *The Annals of Mathematical Statistics*, 28, 678–686 [MR 19:1095e].

61. Graybill, F. A., and Milliken, G. (1969). Quadratic forms and idempotent matrices with random elements. *The Annals of Mathematical Statistics*, 40, 1430–1438 [MR 39:2264].

62. Gundberg, William Roy, Jr. (1984). *A History of Results on Independence of Quadratic Forms in Normal Variates.* MA thesis, Dept. of Mathematics, Arizona State University, Tempe.

63. Gupta, A. K., and Chattopadhyay, A. K. (1979). Gammaization and Wishartness of dependent quadratic forms. *Communications in Statistics–A, Theory and Methods*, 8, 945–951 [MR 81c:62009, Zbl 412.62009].

64. Gupta, A. K., and Varga, T. (1993). *Elliptically Contoured Models in Statistics.* Mathematics and its Applications, vol. 240. Kluwer Academic, Dordrecht [MR 95a:62038].

65. Guttman, Irwin (1982). *Linear Models: An Introduction.* Wiley, New York [MR 84a:62095].

66. Halmos, Paul R. (1951). *Introduction to Hilbert Space and the Theory of Spectral Multiplicity.* Chelsea, New York.

67. Hartwig, Robert E., and Styan, George P. H. (1987). Partially ordered idempotent matrices. In *Proceedings of the Second International Tampere Conference on Statistics: Tampere, Finland, June 1987* (Tarmo Pukkila and Simo Puntanen, eds.), Dept. of Mathematical Sciences, University of Tampere, pp. 361–383.

68. Harville, David A. (1997). *Matrix Algebra from a Statistician's Perspective.* Springer, New York [Zbl 970.46298].

69. Harville, David A., and Kempthorne, Oscar (1997). An alternative way to establish the necessity part of the classical result on the statistical independence of quadratic forms. *Linear Algebra and its Applications*, 264, 205–215.

70. Hocking, Ronald R. (1985). *The Analysis of Linear Models.* Brooks/Cole, Monterey, California [MR 86k:62120, Zbl 625.62054].

71. Hogg, Robert V., and Craig, Allen T. (1958). On the decomposition of certain χ^2 variables. *The Annals of Mathematical Statistics*, 29, 608–610.

72. Hogg, Robert V., and Craig, Allen T. (1978). *Introduction to Mathematical Statistics.* Fourth Edition. Macmillan, New York [MR 57:7817]. [First Edition: 1959, Second: 1965, Third: 1970.]

73. Hotelling, Harold (1944). Note on a matric theorem of A. T. Craig. *The Annals of Mathematical Statistics*, 29, 427–429 [Zbl 060.30805 & 063.02077].

74. Hsu, Pao-Lu (1962). Lecture: A theorem on eigenvalues that is useful in proving the necessity part of the Craig-Sakamoto Theorem for N$(0, I)$. Lemma 2.8.3 in Fang and Zhang [52, pp. 77–79] and Zhang and Fang [247, pp. 123–125]. (See also Zhang, Lu and Qiu [245].)

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

76. Hyakutake, Hiroto, and Siotani, Minoru (1985). Notes on Wishartness and independence of multivariate quadratic forms in correlated normal vectors. *Hiroshima Mathematical Journal*, 15, 75–79 [MR 86m:62100, Zbl 595.62046].

77. Ikeda, Sadao; Hayakawa, Takesi; Hudimoto, Hirosi; Okamoto, Masashi; Siotani, Minoru & Yamamoto, Sumiyasu, eds. (1976). *Essays in Probability and Statistics: A Volume in Honor of Professor Junjiro Ogawa*. Editorial Committee for Publication of Essays in Probability and Statistics, Shinko Tsusho, Tokyo [MR 82g:01054].

78. Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419–426 [MR 25:655, Zbl 136.41103].

79. International Statistical Institute (1985). *The International Statistical Institute (1885–1985)*. Centenary Booklet, iii + 68 pp., Voorburg, The Netherlands.

80. James, G. S. (1952). Notes on a theorem of Cochran. *Proceedings of the Cambridge Philosophical Society*, 48, 443–446 [MR 13:962g, Zbl 046.35806].

81. Jie, Shu Qing, and Li, Pai Chang (1990). Another way to prove Craig's lemma [in Chinese]. *Journal of Central China Normal University, Natural Sciences/Huazhong Shifan Daxue Xuebao, Ziran Kexue Ban (Wuchang)*, 24 (1), 25–26.

82. Johnson, Charles R. (1976). A note on pairs of matrices with product zero. *Journal of Research of the National Bureau of Standards: Section B, Mathematical Sciences*, 80, 333–335.

83. Johnson, Norman L.; Kotz, Samuel, and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Second Edition. Two-volume set. Wiley, New York [Zbl 811.62001, 821.62001]. (Original Edition by Norman L. Johnson and Samuel Kotz: Hougton Mifflin, New York, 1970.

84. Kac, M. (1945). A remark on independence of linear and quadratic forms involving independent Gaussian variables. *The Annals of Mathematical Statistics*, 16, 400–401 [MR 7:310i, Zbl 606.28601].

85. Kamalov, M. K. (1953). Condition for independence of ratios of quadratic forms [in Russian]. *Trudy Instituta Matematiki i Mehaniki im. V.I. Romanovskogo, Akademija Nauk Uzbekstoĭ SSR*, 11, 88–91 [MR 17:871e].

86. Kamalov, M. K. (1955). Moments and cumulants of quadratic and bilinear forms of normally distributed variances [in Russian]. *Trudy Instituta Matematiki i Mehaniki im. V.I. Romanovskogo, Akademija Nauk Uzbekstoĭ SSR*, 15, 69–77 [MR 17:9621].

87. Kawada, Yukiyosi (1950). Independence of quadratic forms in normally correlated variables. *The Annals of Mathematical Statistics*, 21, 614–615 [MR 12:346e, Zbl 041.25905].

88. Kawata, T., and Sakamoto, H. (1949). On the characterization of the normal population by the independence of the sample mean and the sample variance. *Journal of the Mathematical Society Japan*, Vol 1, pp. 111–115.

89. Kendall, Sir Maurice, and Stuart, Alan (1977). *The Advanced Theory of Statistics, Volume 1: Distribution Theory*. Fourth Edition. Griffin, London. (First Edition: 1958; Second: 1963; Third: 1969. Fifth and Sixth Editions by Stuart and Ord [224].)

90. Khatri, C. G. (1959). On conditions for the forms of the type XAX' to be distributed independently or to obey Wishart distribution. *Calcutta Statistical Association Bulletin*, 8, 162–168 [MR 21:4490, Zbl 133.41303].

91. Khatri, C. G. (1961). Cumulants and higher order uncorrelation of certain functions of normal variates. *Calcutta Statistical Association Bulletin*, 10, 93–98 [MR 24:A2466, Zbl 099.35703]. (Errata: 12 (1963), 67 [Zbl 114.10002].)

92. Khatri, C. G. (1961). 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 [MR 26:6189].

93. Khatri, C. G. (1962). Conditions for Wishartness and independence of second degree polynomials in a normal vectors. *The Annals of Mathematical Statistics*, 33, 1002–1007 [MR 25:4594].

94. 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 [MR 29:5342].

95. Khatri, C. G. (1968). Some results for the singular normal multivariate regression models. *Sankhyā Series A*, 30, 267–280 [MR 39:5013, Zbl 177.22802].

96. Khatri, C. G. (1975). Distribution of a quadratic form in normal vectors (multivariate non-central case). In *Satistical Distributions in Scientific Work, Volume I* (G. P. Patil, S. Kotz and J. K. Ord, eds.), Reidel, Dordrecht, Netherlands, pp. 345–354.

97. Khatri, C. G. (1977). Quadratic forms and extension of Cochran's theorem to normal vector variates. In *Multivariate Analysis–IV: Proceedings of the Fourth International Symposium, Dayton, Ohio, 1975*, North-Holland, Amsterdam, pp. 79–94. [MR 56:13448, Zbl 382.62014].

98. Khatri, C. G. (1978). A remark on the necessary and sufficient conditions for a quadratic form to be distributed as chi-squared. *Biometrika*, 65, 239–240 [MR 80a:62066].

99. Khatri, C. G. (1980). Quadratic forms. In *Handbook of Statistics, Volume I* (P. R. Krishnaiah, ed.), North-Holland, Amsterdam, pp. 443–469. [Zbl 578.62051].

100. Khatri, C. G. (1980). The necessary and sufficient conditions for dependent quadratic forms to be distributed as multivariate gamma. *Journal of Multivariate Analysis*, 10, 2, 233–242 [MR 81i:62026, Zbl 431.62028].

101. Khatri, C. G. (1980). Powers of matrices and idempotency. *Linear Algebra and its Applications*, 33, 57–65.

102. Khatri, C. G. (1982). A theorem on quadratic forms for normal variables. In *Statistics and Probability: Essay in Honor of C. R. Rao* (G. Kallianpur, P. R. Krishnaiah and J. K. Ghosh, eds.), North-Holland, Amsterdam, pp. 411–417 [MR 83h:62077, Zbl 487.62042].

103. Klebanov, L. B. (1970). A certain generalization of Craig's theorem [in Russian]. *Doklady Akademii Nauk SSSR*, 195, 763–764 [MR 43:1322, Zbl 215.54303]. [Translated into English: *Soviet Mathematics Doklady*, 11, 1551–1552 (1971).]

104. Knight, John L. (1985). The joint characteristic function of linear and quadratic forms of non-normal variables. *Sankhyā Series A*, 47, 231–238 [MR 87k:60050, Zbl 581.60018].

105. Koerts, Johan, and Abrahamse, A. P. J. (1969). *On the Theory and Application of the General Linear Model*. Rotterdam University Press [MR 51:7170, Zbl 225.62081].

106. Kolmogoroff, A. (1933). *Grundbegriffe der Wahscheinlichkeitsrechnung* [in German]. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 2, no. 3, Julius Springer, Berlin [Zbl 007.21601]. (Cf. Section 5 "Unabhängigkeit", pp. 8–11; footnote, p. 10.)

107. Koopmans, T. (1942). Serial correlation and quadratic forms in normal variables. *The Annals of Mathematical Statistics*, 13, 14–33 [MR 4:22e].

108. Krafft, Olaf (1978). *Lineare statistische Modelle und optimale Versuchspläne* [in German]. Studia Mathematica/Mathematische Lehrbücher, Band 30. Vandenhoeck & Ruprecht, Göttingen [MR 80c:62092].

109. Kshirsagar, Anant M. (1983). *A Course in Linear models*. Statistics: Textbooks and Monographs, vol. 45. Marcel Dekker, New York [MR 85c:62183].

110. Laha, R. G. (1956). On the stochastic independence of two-degree polynomial statistics in normally distributed variates. *The Annals of Mathematical Statistics*, 27, 790–796 [MR 18:160c, Zbl 073.35603].

111. Laha, R. G., and Lukacs, E. (1960). On certain functions of normal variates which are uncorrelated of a higher order. *Biometrika*, 47, 175–176 [MR 22:1950, Zbl 094.33502].

112. Laha, R. G., and Lukacs, E. (1964). *Applications of Characteristic Functions*. Griffin's Statistical Monographs and Courses, no. 14. Griffin, London [MR 30:2582, Zbl 117.36402].

113. Lancaster, H. O. (1954). Traces and cumulants of quadratic forms in normal variables. *Journal of the Royal Statistical Society, Series B*, 16, 247–254 [MR 16:1128h, Zbl 058.12001].

114. Lancaster, H. O. (1969). *The Chi-Squared Distribution*. Wiley, New York [MR 40:6667, Zbl 193.17802].

115. Leipnik, R. B. (1958). Moment generating functions of quadratic forms in serially correlated normal variables. *Biometrika*, 45, 198–210 [MR 19:1095a].

116. Letac, G., and Massam, H. (1995). Craig-Sakamoto's theorem for the Wishart distributions on symmetric cones. *Annals of the Institute of Statistical Mathematics Tokyo*, 47, 785–799 [MR 97g:62091, Zbl 847.62042].

117. Li, Chi-Kwong (2000). A simple proof of the Craig-Sakamoto theorem. *Linear Algebra and its Applications*, Eighth Special Issue on Linear Algebra and Statistics, in press.

118. Li, Chi-Kwong, and Styan, George P. H. (1999). An extension of the Craig-Sakamoto theorem. Preprint, 4 pp.

119. Lin, Chun-tu (1984). Extrema of quadratic forms and statistical applications. *Communications in Statistics–Theory and Methods*, 13, 1517–1520 [Zbl 554.62055].

120. Liu, Wan Rong (1991). Necessary and sufficient conditions for mutual independence of random matrices of quadratic forms and their applications [in Chinese]. *Hunan Shifan Daxue Ziran Kexue Xuebao/ Acta Scientiarum Naturalium Universitatis Normalis Hunanensis (Changsha)*, 14 (3), 209–213 [Zbl 737.62048].

121. Loynes, R. M. (1966). On idempotent matrices. *The Annals of Mathematical Statistics*, 37, 295–296 [MR 32:1209].

122. Lukacs, Eugene (1952). The stochastic independence of symmetric and homogeneous linear and quadratic statistics. *The Annals of Mathematical Statistics*, 23, 442–449 [MR14:297e, Zbl 047.38102].

123. Lukacs, Eugene (1970). *Characteristic Functions*. Second Edition (revised and enlarged). Griffin, London [MR 49:11595, Zbl 201.20404].

124. Luther, N. Y. (1965). Decomposition of symmetric matrices and distributions of quadratic forms. *The Annals of Mathematical Statistics*, 36, 683–690 [MR 30:2592].

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

126. Madow, William G. (1940). The distribution of quadratic forms in non-central normal random variables. *The Annals of Mathematical Statistics*, 11, 100–103 [MR 1:248f].

127. Magnus, Jan R. (1979). The expectation of products of quadratic forms in normal variables: the practice. *Statistica Neerlandica*, 33, 131–136 [MR 51k:62024, Zbl 412.60024].

128. Magnus, Jan R. (1988). *Linear Structures*. Griffin's Statistical Monographs & Courses, no. 42. Griffin, London [MR 89j:15003].

129. Magnus, Jan R., and Neudecker, Heinz (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Revised Edition. Wiley, Chichester [Zbl 990.08016]. (Original edition: 1988.)

130. Mäkeläinen, Timo (1966) On quadratic forms in normal variables. *Commentationes Physico-Mathematicae, Societas Scientiarum Fennica (Helsinki)*, 31 (12), 1-6 [MR 33:8053].

131. Mäkeläinen, Timo, and Styan, George P. H. (1976). A decomposition of an idempotent matrix where nonnegativity implies idempotence and none of the matrices need be symmetric. *Sankhyā Series A* 38, 400–403 [MR 57:12554, Zbl 412:15018].

132. Marcus, Marvin (1998). On a determinant result of I. Olkin. *Linear Algebra and its Applications*, 277, 237–238 [MR 99b:15007].

133. Marsaglia, George, and Styan, George P. H. (1974). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra*, 2, 269–292 [MR 52:5711, Zbl 297.15003].

134. Matérn, Bertil (1949). Independence of non-negative quadratic forms in normally correlated variables. *The Annals of Mathematical Statistics*, 20, 119–120. [MR 10:553b, Zbl 041.25904].

135. Mathai, A. M., and Provost, Serge B. (1992). *Quadratic Forms in Random Variables: Theory and Applications*. Statistics: Textbooks and Monographs, vol. 126. Marcel Dekker, New York [MR 94g:62110, Zbl 792.62045].

136. Mathai, A. M.; Provost, Serge B., and Hayakawa, Takesi (1995). *Bilinear Forms and Zonal Polynomials*. Lecture Notes in Statistics, vol. 102. Springer-Verlag, New York [MR 98k:62016, Zbl 832.62044].

137. Mathew, Thomas, and Nordström, Kenneth (1997). Wishart and chi-square distributions associated with matrix quadratic forms. *Journal of Multivariate Analysis*, 61, 129–143 [MR 98k:62094, Zbl 970.48478].

138. Matusita, Kameo (1949). Note on the independence of certain statistics. *Annals of the Institute of Statistical Mathematics Tokyo*, 1, 79–82 [MR 11:260e, Zbl 045.41103].

139. Matusita, Kameo (1992). Institute of Statistical Mathematics Tokyo. *The IMS Bulletin*, 21, 108–109.

140. McCloskey, Joseph P. (1984). A unified approach to the statistical distribution of complex quadratic form. *Industrial Mathematics*, 34, 29–52 [MR 86j:62119, Zbl 544.62051].

141. McCloskey, J. P. (1985). Properties of normal, r-potent matrices. *Linear and Multilinear Algebra*, 17, 325–336 [MR 86m:15010].

142. Mirsky, L. (1990). *An Introduction to Linear Algebra.* Second Corrected Reprint Edition. Dover, New York [MR 91j:15001]. (Original Edition: 1955, Oxford University Press.)

143. Mitra, S. K. (1968). A new class of g-inverse of square matrices. *Sankhyā Series A*, 30, 323–330 [MR 39:4180, Zbl 232.62025].

144. Mitra, S. K., and 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 [MR 45:1938].

145. Mitra, Sujit K., and Rao, C. Radhakrishna (1968). Simultaneous reduction of a pair of quadratic forms. *Sankhyā Series A*, 39, 313–322 [MR 39:2784, Zbl 198.35204].

146. Mitra, S. K., and Rao, C. Radhakrishna (1972). Generalized inverse of a matrix and its applications. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (University of California, Berkeley, California, 1970/1971)*, University of California Press, Berkeley, California, Vol. I: Theory of Statistics, pp. 601–620 [MR 53:6906].

147. Moser. Barry Kurt (1996). *Linear Models: A Mean Model Approach.* Academic Press, San Diego [MR 98d:62121].

148. Motzin. T. S., and Taussky, Olga (1952). Pairs of matrices with property *L*. *Transactions of the American Mathematical Society*, 73. 108–114.

149. Nagase, G. (1973). On the covariance between two quadratic forms in a general set up. *Communications in Statistics*, 2, 281–284 [MR 48:3185, 267.62006].

150. Nagase, G., and Banerjee, K. S. (1973). A note on second degree polynomial statistics. *Communications in Statistics*, 1, 525-531. [MR 50:1376].

151. Nagase, G., and Banerjee, K. S. (1975). A note on independence of quadratic forms in noncentral normal variates. *Calcutta Statistical Association Bulletin*, 24, 93–96, 131–133 [MR 55:4516, Zbl 364.62054].

152. Nagase, G., and Banerjee, K. S. (1976). Independence of second degree polynomial statistics in normal variates [Letter to the Editor]. *The American Statistician*, 30, 201 [MR 56:13449, Zbl 333.62033].

153. Nelder, J. A. (1951). A note on the statistical independence of quadratic forms in the analysis of variance. *Biometrika*, 38, 482–483 [MR 13:668b, Zbl 043.13607].

154. Neudecker, Heinz (1985). On the dispersion matrix of a matrix quadratic form connected with the noncentral Wishart distribution. *Linear Algebra and its Applications*, 70, 257–262 [MR 86m:62017, Zbl 569.60025].

155. Neudecker, H. (1990). The variance matrix of a matrix quadratic form under normality assumptions: a derivation based on its moment-generating function. *Statistics*, 21, 455–459.

156. Nordström, Kenneth (1990). Some further aspects of the Löwner-ordering antitonicity of the Moore-Penrose inverse. *Communications in Statistics–Theory and Methods*, 18, 4471–4489 [MR 91c:15002].

157. Ogasawara, Tôzirô, and Takahashi, Masayuki (1951). Independence of quadratic quantities in a normal system. *Journal of Science of the Hiroshima University Series A (Mathematics, Physics, Chemistry)*, 15, 1–9 [MR 13:142d, Zbl 045.41102].

158. Ogawa, Junjiro (1946). 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.

159. Ogawa, Junjiro (1946). On the independence of statistics of quadratic forms [in Japanese]. *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 2, 98–111. (Translated as [160].)

160. Ogawa, Junjiro (1947). On the independence of statistics of quadratic forms [in English]. *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 3, 137–151. (Translation of [159].)

161. Ogawa, Junjiro (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. (Translated as [162].)

162. Ogawa, Junjiro (1949). On the independence of bilinear and quadratic forms of a random sample from a normal population. *Annals of the Institute of Statistical Mathematics Tokyo*, 1, 83–108 [MR 11:260f, Zbl 041.40102]. (Translation of [161].)

163. Ogawa, Junjiro (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. (Translated as [164].)

164. Ogawa, Junjiro (1950). Independence of statistics of quadratic forms for non-central normal distribution [in English]. *Osaka Mathematical Journal*, 2, 151-159 [MR 12,509h, Zbl 040.36305]. (Translation of [163].)

165. Ogawa, Junjiro (1993). A history of the development of Craig-Sakamoto's theorem viewed from Japanese standpoint [in Japanese, with English summary]. *Proceedings of the Institute of Statistical Mathematics Tokyo*, 41, 47–59.

166. Ogawa, Junjiro, and Olkin, Ingram (1997). A tale of two countries: The Craig-Sakamoto-Matusita Theorem. Revised preprint, July 1, 1997, 20 pp. (Original version: September 1, 1995.)

167. Oktaba, Wiktor, and Kieloch, Andrzej (1993). Wishart distributions in the multivariate Gauss-Markoff model with singular covariance matrix. *Applications of Mathematics (Prague)*, 38, 61–66.

168. Olkin, Ingram (1957). An exposition on the independence of quadratic, bilinear and linear forms for normal variates. With addenda. Preprint, 15 pp.

169. Olkin, Ingram (1997). A determinantal proof of the Craig-Sakamoto theorem. *Linear Algebra and its Applications*, 264, 217–223. [MR 98h:15012, Zbl 887.15009].

170. Olkin, Ingram (1998). Why is matrix analysis part of the statistics curriculum? *Student (Neuchâtel)*, 2, 343–348.

171. Pavur, R. J., and Lewis, T. O. (1986). The role of powers of matrices in determining the distribution of quadratic forms involving the complex normal distribution. *Linear Algebra and its Applications*, 76, 71–84 [MR 87e:15031].

172. Plackett, R. L. (1960). *Principles of Regression Analysis*. Oxford University Press [MR 28:2604, Zbl 091.31301].

173. Poltronieri, J. (1984). Formas cuadraticas vectoriales y formas lineales vectoriales [in Spanish]. *Ciencia y Tecnología (Editorial Universidad de San José, Costa Rica)*, 8, 153–165 [MR 87d:62104, Zbl 609.62073].

174. Pringle, R. M., and Rayner, A. A. (1971). *Generalized Inverse Matrices with Applications to Statistics* Griffin's Statistical Monographs & Courses, no. 28. Griffin, London [MR 47:3410, Zbl 231.15008].

175. Provost, Serge B. (1994). On the Craig-Sakamoto lemma. *Pakistan Journal of Statistics*, 10, 205–212 [Zbl 831.62037].

176. Provost, Serge B. (1996). On Craig's theorem and its generalizations. *Journal of Statistical Planning and Inference*, 53, 311–321 [MR 97e:62090, Zbl 960.46936].

177. Puntanen, Simo, and Styan, George P. H. (1996). An equivalence relation for two symmetric idempotent matrices: First Solution to Problem 95.3.3 (proposed by Shuangzhe Liu and Wolfgang Polasek). *Econometric Theory*, 12, 590–591. (See also [246].)

178. Puntanen, Simo; Styan, George P. H., and Zhang, Fuzhen (1998). Properties of idempotent matrix: Solution to Problem 97.3.3 (proposed by Robert E. Hartwig and Götz Trenkler) in *Econometric Theory*. In Report A322, Dept. of Mathematics, Statistics & Philosophy, University of Tampere, Tampere, Finland, pp. 7–8.

179. Puntanen, Simo; Styan, George P. H., and Zhang, Fuzhen (1998). Order invariability of idempotent matrix: Solution to Problem 97.3.4 (proposed by K. Han, E. Im and M. S. Snow), in *Econometric Theory*. In Report A322, Dept. of Mathematics, Statistics & Philosophy, University of Tampere, Tampere, Finland, pp. 10–12.

180. Qin, Huaizhen; Cui, Hengjian, and Li, Yong (1998). The n.s. conditions for the ratio of two quadratic forms to have an F-distribution and its applications. *Communications in Statistics–Theory and Methods*, 27, 453–471 [Zbl 980.21554].

181. Rao, C. Radhakrishna (1962). A note on a generalized inverse of a matrix with application to problems in mathematical statistics. *Journal of the Royal Statistical Society Series B*, 24, 152–158 [MR 25:1596, Zbl 121.14502]. (Reprinted in [38], pp. 316–324.)

182. Rao, C. Radhakrishna (1966). Generalized inverse for matrices and its applications in mathematical statistics. In *Research Papers in Statistics: Festschrift for J. Neyman* (F. N. David, ed.), Wiley, London, pp. 263–279 [MR 35:198, Zbl 203.21005].

183. Rao, C. Radhakrishna (1973). *Linear Statistical Inference and Its Applications* Second Edition. Wiley, New York [MR 49:11677].

184. Rao, C. Radhakrishna (1978). Least squares theory for possibly singular models. *The Canadian Journal of Statistics*, 6, 19–23 [MR 81a:62064].

185. Rao, C. Radhakrishna, and Mitra, Sujit Kumar (1971). *Generalized Inverse of Matrices and Its Applications*. Wiley, New York [MR 49:2780].

186. Rao, C. Radhakrishna, and Rao, M. Bhaskara (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics*. World Scientific, Singapore.

187. Rao, C. Radhakrishna, and Toutenburg, Helge (1999). Second Edition. *Linear Models: Least Squares and Alternatives*. Springer, New York. (Original Edition: 1995.)

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

189. Rayner, A. A. (1974). Quadratic forms and degeneracy: a personal view. Invited lecture, Annual Conference, South African Statistical Association, Pietermaritzburg.

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

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

192. Reid, John G., and Driscoll, Michael F. (1988). An accessible proof of Craig's theorem in the noncentral case. *The American Statistician*, 42, 139–142.

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

194. Ruben, Harold (1978). On quadratic forms and normality. *Sankhyā Series A*, 40, 156–173 [MR 80i:62010, Zbl 409.60019].

195. Rudin, Walter (1987). *Real and Complex Analysis*. Third Edition. McGraw-Hill, New York [Zbl 925.00005]. (Original Edition: 1966 [Zbl 142.01701]; Second Edition: 1974 [Zbl 278.26001].)

196. Sakamoto, Heihachi (1944). 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.

197. Sakamoto, Heihachi (1944). On the independence of two statistics [in Japanese]. *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 1 (9), 1–25.

198. Sakamoto, Heihachi (1946). On the independence of statistics and the judgement of degrees of freedom in the theory of analysis of variance [in Japanese]. *Research Memoirs of the Institute of Statistical Mathematics Tokyo*, 2, 155-162 [Zbl 041.46703]. (Translated as [200].)

199. Sakamoto, Heihachi (1948). On the independence and degrees of freedom of the maximum likelihood estimates [in Japanese]. *Bulletin of Mathematical Statistics*, 2 (1), 36–37 & 2 (2) 24–27.

200. Sakamoto, Heihachi (1949). On the independence of statistics and the judgement of degrees of freedom in the theory of analysis of variance [in English]. *Annals of the Institute of Statistical Mathematics Tokyo*, 1, 109–122 [MR 11:260g, Zbl 045.41201]. (Translation of [198].)

201. Sakamoto, Heihachi (1949). On independence of statistical quantities [in Japanese]. *Sûgaku*, 1, 263–274 [MR 13:366f].

202. Satake, Ichiô (1975). *Linear Algebra* [in English]. Translated from the Japanese and Revised by Sebastian Koh, Tadatoshi Akiba and Shin-ichiro Ihara. Pure and Applied Mathematics, vol. 29. Marcel Dekker, New York [MR 53:5616, Zbl 302.15001].

203. Scarowsky, Issie (1973). *Quadratic Forms in Normal Variables.* MSc thesis, Dept. of Mathematics, McGill University, Montréal.

204. Scarowsky, Issie, and Styan, George P. H. (1982). Some notes on quadratic forms in normal variables. Report No. 82-14, Dept. of Mathematics, McGill University, Montréal.

205. Schall, Robert, and Dunne, Timothy T. (1987). A note on the chi-squaredness of quadratic forms. *Sankhyā Series A*, 49, 415–418. [MR 91b:62098, Zbl 641.62034].

206. Searle, Shayle R. (1966). *Matrix Algebra for the Biological Sciences (Including Applications in Statistics).* Wiley, New York [Zbl 137.39702].

207. Searle, Shayle R. (1982). *Matrix Algebra Useful for Statistics.* Wiley, New York [MR 83i:15002].

208. Searle, S. R. (1984). Detailed proofs (class notes) of necessary and sufficient conditions for independence and χ^2-distribution properties of quadratic forms in normal variables. Paper No. BU-834-M, Biometrics Unit, Cornell University, Ithaca, New York, 19 pp.

209. Searle, Shayle R. (1997). *Linear Models.* Wiley Classics Library Edition. Wiley, New York. (Original version: 1971 [MR 45:2868].)

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

211. Seber, George A. F. (1973). Independence and normal variables. *The American Statistician*, 27, 110.

212. Seber, George A. F. (1977). *Linear Regression Analysis.* Wiley, New York [MR 55:9428].

213. Seber, George A. F. (1980). *The Linear Hypothesis: A General Theory.* Second Edition. Griffin's Statistical Monographs & Courses, no. 19. Griffin, London.

214. Šemrl, Peter (1996). On a matrix version of Cochran's statistical theorem. *Linear Algebra and its Applications*, 237/238, 477–487 [MR 97b:62087, Zbl 847.62043].

215. Shah, B. K. (1963). Distribution of definite and indefinite quadratic forms from a non-central normal distribution. *The Annals of Mathematical Statistics*, 34, 186–190 [MR 26:1953, Zbl 246.62028]. (Correction: 39, 289 (1968) [MR 36:3431].)

216. Shah, B. K. (1970). Distribution theory of a positive definite quadratic form with matrix argument. *The Annals of Mathematical Statistics*, 41, 692–697 [MR 41:4709].

217. Shah, B. K., and Khatri C. G. (1961). Distribution of a definite quadratic form for non-central normal variates. *The Annals of Mathematical Statistics*, 32, 883–887 [MR 32:A4196].

218. Shanbhag, D. N. (1966). On the independence of quadratic forms. *Journal of the Royal Statistical Society Series B*, 28, 582–583. [MR 34:8532, Zbl 144.41605].

219. Shanbhag, D. N. (1968). Some remarks concerning Khatri's result on quadratic forms. *Biometrika*, 55, 593–596 [Zbl 177.47403].

220. Shanbhag, D. N. (1970). On the distribution on a quadratic form. *Biometrika*, 57, 222–223 [Zbl 193.18101].

221. Sjögren, Peter (1968). Ett matrislemma till Cochrans stas [in Swedish]. *Nordisk Matematisk Tidskrift (Oslo)*, 16, 157–158 [MR 39:2786].

222. Smiley, M. F. (1964). The spectral theorem for finite matrices and Cochran's theorem. *The Annals of Mathematical Statistics*, 35, 443–444 [MR 28:1687].

223. Srivastava, Muni S., and Khatri, C. G. (1979). *An Introduction to Multivariate Statistics.* North-Holland, Amsterdam [MR 53:1878].

224. Stuart, Alan, and Ord, J. Keith (1994). *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory.* Sixth Edition. Edward Arnold, London & Halsted Press, New York [Zbl 880.62012]. (Fifth Edition: 1991 [Zbl 621.62001]. First–Fourth Editions by Kendall and Stuart [89].)

225. Styan, George P. H. (1970). Notes on the distribution of quadratic forms in singular normal variables. *Biometrika*, 57, 567–572 [Zbl 264:62006].

226. Styan, George P. H., ed. (1990). *The Collected Papers of T. W. Anderson: 1943–1985.* Two-volume set. Wiley, New York [MR 91j: 01064].

227. Styan, George P. H., ed. (1991). Allen Thornton Craig: 1904–1978 & The Allen T. Craig Lectures at the University of Iowa. *The IMS Bulletin*, 20, 389–390.

228. Styan, George P. H., and Takemura, Akimichi (1983). Rank additivity and matrix polynomials. In *Studies in Econometrics, Time Series and Multivariate Statistics in Honor of Theodore W. Anderson* (Samuel Karlin, Takeshi Amemiya and Leo A. Goodman, eds.), Academic Press, New York, pp. 545–558 [MR 85f:62004, 86e:15003; Zbl 586:15002].

229. Tan, W. Y. (1975). Some matrix results and extensions of Cochran's theorem. *SIAM Journal on Applied Mathematics*, 28, 547–554 [MR 51:2159, Zbl 296.62039]. (Errata: 30 (1976), 608–610 [MR 53:14794, Zbl 351.62011].)

230. Tan, W. Y. (1977). On the distribution of quadratic forms in normal variables. *The Canadian Journal of Statistics*, 5, 241–250 [MR 58:18897, Zbl 379.62038].

231. Taussky, Olga (1958). On a matrix theorem of A. T. Craig and H. Hotelling. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings, Series A*, vol. 61, & *Indagationes Mathematicae*, vol. 20, pp. 139–141 [Zbl 083.00602].

232. Turin, G. L. (1960). The characteristic function of Hermitian quadratic forms in complex normal variables. *Biometrika*, 47, 199–201 [MR 22:8546, Zbl 087.33802].

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

234. Wang, Tonghui (1997). Versions of Cochran's theorem for general quadratic expressions in normal matrices. *Journal of Statistical Planning and Inference*, 58, 283–297 [MR 98h:62097, Zbl 970.37873].

235. Wang, Tonghui (1997). Wishartness of quadratic forms under elliptically contoured settings. *Publications de l'Institut de Statistique de l'Université de Paris*, 3, 31–45.

236. Wang, Tonghui; Fu, Ronglin, and Wong, Chi Song (1996). Cochran theorems for multivariate components of variance models. *Sankhyā, Series A*, 58, 328–342.

237. Wang, Tonghui, and Wong, Chi Song (1995). Cochran theorems for a multivariate elliptically contoured model. *Journal of Statistical Planning and Inference*, 43, 257–270 [MR 95m:62116, Zbl 812.62055].

238. Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York [MR 26:1949, Zbl 173.45805].

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

240. Wong, Chi Song; Masaro, Joe, and Wang, Tonghui (1991). Multivariate versions of Cochran's theorems. *Journal of Multivariate Analysis*, 39, 154–174 [MR 92i:62104, Zbl 749.62037].

241. Wong, Chi Song, and Wang, Tonghui (1993). Multivariate versions of Cochran's theorems II. *Journal of Multivariate Analysis*, 44, 146–159 [MR 94b:62044, Zbl 765.62057].

242. Wong, Chi Song, and Wang, Tonghui (1995). Laplace-Wishart distributions and Cochran theorems. *Sankhyā Series A*, 57, 342–359 [MR 97d:62106, Zbl 970.05835].

243. Young, Dean M.; Odell, Patrick L., and Seaman, John W., Jr. (1994). On the independence of the sample mean and translation-invariant statistics for matrix normal distributions. *The Australian Journal of Statistics*, 36, 95–100 [Zbl 829.62058].

244. Zelen, M. (1962). Linear estimation and related topics. In *Survey of Numerical Analysis*, McGraw-Hill, New York, pp. 558–584 [MR 24:3747].

245. Zhang, Bao-Xue; Lu, Chang-Yu, and Qiu, Hong-Bin (1998). The Hsu's proof of the necessity part of the classical result on the statistical independence of quadratic forms. Dept. of Mathematics, Northeast Normal University, Changchun, China. Preprint, 7 pp.

246. Zhang, Fuzhen, and Styan, George P. H. (1996). An equivalence relation for two symmetric idempotent matrices: Second Solution to Problem 95.3.3 (proposed by Shuangzhe Liu and Wolfgang Polasek). *Econometric Theory*, 12, 591–592. (See also [177].)

247. Zhang, Yao-Ting, and Fang, Kai-Tai (1983). *Introduction to Multivariate Statistical Analysis* [in Chinese]. Academic Press, Beijing. (Expanded and translated into English as Fang & Zhang (1990) [52]. Cf. Lemma 2.8.3, pp. 123–125.)

248. Zielinski, A. (1985). A note on Nabeya's proof of the Sakamoto-Craig lemma. *Statistical Decisions*, Supplement Issue no. 2, 299–300.