REVSTAT – Statistical Journal
Volume 16, Number 3, July 2018, 315–322

ON THE DISTRIBUTION OF A QUADRATIC FORM IN NORMAL VARIATES

Author: Jin Zhang
– School of Mathematics and Statistics, Yunnan University,
Kunming, Yunnan, 650091, China
jinzhang@ynu.edu.cn

Received: August 2016 Revised: September 2016 Accepted: September 2016

Abstract:

- It is a well-known theorem in linear models that the idempotency of a matrix is a sufficient and necessary condition for a quadratic form in normal variates to have a chi-square distribution, but its proofs in the early literature were incorrect or incomplete. Driscoll (1999) provided an improved proof, and this article presents a simple proof. More importantly, we establish and prove a generalized theorem.

Key-Words:

- *eigenvalues; idempotent matrices; moment-generating function; normality.*

AMS Subject Classification:

- 62J05, 62H10, 62E15.

1. INTRODUCTION

There is a rich literature on the distribution and independence of quadratic forms in normal random vectors (e.g, Rao, 1973; Graybill, 1976; Driscoll and Gundberg, 1986; Mathai and Provost, 1992; Jorgensen, 1993; Driscoll, 1999; Christensen, 2002; Ravishanker and Dey, 2002; Ogawa and Olkin, 2008), which play an important role in linear models and multivariate statistical analysis.

Let $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote the k-dimensional normal distribution with mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$, and let $\chi^2_m(\lambda)$ be the noncentral chi-square distribution with m degrees of freedom and noncentrality parameter λ. The two well-known theorems below establish sufficient and necessary conditions for the independence and distributions of quadratic forms in normal variates.

Theorem 1. *Let $\boldsymbol{x} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > \mathbf{0}$, and $\boldsymbol{A}$ and $\boldsymbol{B}$ be $k \times k$ real symmetric matrices. Then $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$ are independently distributed if and only if $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B} = \mathbf{0}$.*

Theorem 2. *Let $\boldsymbol{x} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > \mathbf{0}$, and $\boldsymbol{A}$ be a $k \times k$ real symmetric matrix. Then $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \sim \chi^2_m(\lambda)$ with $\lambda = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ if and only if $\boldsymbol{A}\boldsymbol{\Sigma}$ is idempotent of rank m.*

Unfortunately, the proofs of the two theorems in the early literature are incorrect, incomplete or misleading, especially for Theorem 1 (Driscoll and Gundberg, 1986; Driscoll, 1999; Ogawa and Olkin, 2008). Thus, many improved proofs for Theorem 1 have been obtained by Reid and Driscoll (1988), Driscoll and Krasnicka (1995), Letac and Massam (1995), Provost (1996), Olkin (1997), Marcus (1998), Li (2000), Matsuura (2003), Ogawa and Olkin (2008), Carrieu and Lassère (2009), Carrieu (2010), Bonnefond (2012), Zhang and Yi (2012), and many others. However, there is only one improved proof of Theorem 2 given by Driscoll (1999). In addition, Liu *et al.* (2009) and Duchesne and Lafaye De Micheaux (2010) discussed the computational issues in Theorem 2.

A simple proof of Theorem 2 is presented in Section 2, using elementary calculus and matrix algebra. We give a counter example of Theorem 2 in Section 3, where $\boldsymbol{\Sigma}$ is singular. Then we establish and prove its extension in Theorem 3 for the general case, where $\boldsymbol{\Sigma}$ can be singular or nonsingular.

2. A SIMPLE PROOF OF THEOREM 2

The proof of sufficiency for Theorem 2 is quite easy, but showing necessity is difficult. In fact, its proofs in the early literature were incorrect or incomplete, according to Driscoll (1999), who provided an improved proof of Theorem 2, based on the moment-generating function and cumulants. We now present a simple proof of Theorem 2, using the moment-generating function of $\chi^2_m(\lambda)$:

$$M_{\chi^2_m(\lambda)}(t) = (1-2t)^{-\frac{m}{2}} e^{\frac{2t\lambda}{1-2t}}, \quad t < 1/2.$$

Proof of Theorem 2:

Sufficiency. Suppose $(\boldsymbol{A\Sigma})^2 = \boldsymbol{A\Sigma}$ and $r(\boldsymbol{A\Sigma}) = m$, where $\boldsymbol{\Sigma} = \boldsymbol{BB}'$ and $\tilde{\boldsymbol{A}} = \boldsymbol{B}'\boldsymbol{AB}$. Then $\tilde{\boldsymbol{A}}^2 = \tilde{\boldsymbol{A}}$ and $r(\tilde{\boldsymbol{A}}) = m$. Thus, there exists an orthogonal matrix $\boldsymbol{P}$ such that

$$\tilde{\boldsymbol{A}} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{P}' = \boldsymbol{P}_1\boldsymbol{P}_1',$$

where $\boldsymbol{P} = (\boldsymbol{P}_1, \boldsymbol{P}_2)$, $\boldsymbol{P}_1'\boldsymbol{P}_1 = \boldsymbol{I}_m$ and $\boldsymbol{z} = \boldsymbol{P}_1'\boldsymbol{B}^{-1}\boldsymbol{x} \sim N_m(\boldsymbol{P}_1'\boldsymbol{B}^{-1}\boldsymbol{\mu},\ \boldsymbol{I})$. It follows that

$$\boldsymbol{x}'\boldsymbol{Ax} = \boldsymbol{z}'\boldsymbol{z} \sim \chi^2_m(\lambda),$$

where $\lambda = \frac{1}{2}(\boldsymbol{P}_1'\boldsymbol{B}^{-1}\boldsymbol{\mu})'(\boldsymbol{P}_1'\boldsymbol{B}^{-1}\boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A\mu}$.

Necessity. Suppose $\boldsymbol{x}'\boldsymbol{Ax} \sim \chi^2_m(\lambda)$. Let $\boldsymbol{P} = (\boldsymbol{p}_1, \cdots, \boldsymbol{p}_k)$ be an orthogonal matrix such that $\boldsymbol{P}'\tilde{\boldsymbol{A}}\boldsymbol{P} = \boldsymbol{\Lambda} = \text{diag}(\lambda_1, ..., \lambda_k)$, where $\lambda_1 \geq \cdots \geq \lambda_k$ are eigenvalues of $\tilde{\boldsymbol{A}}$. Then

$$\boldsymbol{x}'\boldsymbol{Ax} = \boldsymbol{z}'\boldsymbol{\Lambda z} = \sum_{i=1}^k \lambda_i z_i^2, \quad M_{\boldsymbol{x}'\boldsymbol{Ax}}(t) = \prod_{i=1}^k M_{z_i^2}(t\lambda_i),$$

where $\boldsymbol{z} = \boldsymbol{P}'\boldsymbol{B}^{-1}\boldsymbol{x} \sim N_k(\boldsymbol{P}'\boldsymbol{B}^{-1}\boldsymbol{\mu},\ \boldsymbol{I})$ and $z_1, ..., z_k$ are independent. Hence,

$$(1-2t)^{-\frac{m}{2}} e^{\frac{2t\lambda}{1-2t}} = \prod_{i=1}^k (1-2t\lambda_i)^{-\frac{1}{2}} e^{\frac{t\lambda_i}{1-2t\lambda_i}(\boldsymbol{p}_i'\boldsymbol{B}^{-1}\boldsymbol{\mu})^2}$$

for $t < 1/2$ and $t\lambda_i < 1/2$ $(i = 1, ..., k)$. Comparing the discontinuous points of the two functions on both sides results in

$$(1-2t)^{-\frac{m}{2}} = \prod_{i=1}^k (1-2t\lambda_i)^{-\frac{1}{2}} = |\boldsymbol{I}_k - 2t\tilde{\boldsymbol{A}}|^{-\frac{1}{2}},$$

which implies that $\lambda_1 = \cdots = \lambda_m = 1$ and $\lambda_{m+1} = \cdots = \lambda_k = 0$.

Thus, $\tilde{\boldsymbol{A}}$ or $\boldsymbol{A\Sigma}$ is idempotent of rank m. The proof is completed. □

3. DISTRIBUTIONS OF QUADRATIC FORMS IN THE GENERAL CASE

We now discuss the distribution of quadratic form $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ in the general case, where $\boldsymbol{x} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ but $\boldsymbol{\Sigma}$ can be singular or not. First, it should be pointed out that Theorem 2 is not true when $\boldsymbol{\Sigma}$ is singular. Below is a counter example.

Let $\boldsymbol{A} = \boldsymbol{I}_2$ and $\boldsymbol{x} = (z, 1)'$, where $z \sim N(0, 1)$. Then $\boldsymbol{x} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

It is clear that $(\boldsymbol{A}\boldsymbol{\Sigma})^2 = \boldsymbol{A}\boldsymbol{\Sigma}$ and its rank $r(\boldsymbol{A}\boldsymbol{\Sigma}) = 1$, but

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = z^2 + 1,$$

the distribution of which is not $\chi_1^2(\lambda)$ with $\lambda = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu} = \frac{1}{2}$.

To generalize Theorem 2, we have the following Theorem 3, which reduces to Theorem 2 if $\boldsymbol{\Sigma}$ is nonsingular. The proof for Theorem 3 is based on the moment-generating function of quadratic function $Q = \boldsymbol{z}'\boldsymbol{A}\boldsymbol{z} + \boldsymbol{b}'\boldsymbol{z} + c$:

$$M_Q(t) = |\boldsymbol{I} - 2t\boldsymbol{A}|^{-\frac{1}{2}} e^{ct + \frac{t^2}{2}\boldsymbol{b}'(\boldsymbol{I}-2t\boldsymbol{A})^{-1}\boldsymbol{b}}$$

for small $|t|$ such that $\boldsymbol{I} - 2t\boldsymbol{A} > \boldsymbol{0}$, where $\boldsymbol{z} \sim N_k(\boldsymbol{0},\ \boldsymbol{I})$, $\boldsymbol{A}$ is a real symmetric matrix, $\boldsymbol{b}$ is a k-dimensional real vector, and c is a real number. In fact,

$$\begin{aligned} M_Q(t) &= \int (2\pi)^{-\frac{k}{2}} e^{t(\boldsymbol{z}'\boldsymbol{A}\boldsymbol{z}+\boldsymbol{b}'\boldsymbol{z}+c)-\frac{1}{2}\boldsymbol{z}'\boldsymbol{z}} d\boldsymbol{z} \\ &= e^{ct+\frac{t^2}{2}\boldsymbol{b}'\boldsymbol{A}_t\boldsymbol{b}} \int (2\pi)^{-\frac{k}{2}} e^{-\frac{1}{2}(\boldsymbol{z}-t\boldsymbol{A}_t\boldsymbol{b})'\boldsymbol{A}_t^{-1}(\boldsymbol{z}-t\boldsymbol{A}_t\boldsymbol{b})} d\boldsymbol{z}, \end{aligned}$$

where $\boldsymbol{A}_t = (\boldsymbol{I} - 2t\boldsymbol{A})^{-1}$.

Theorem 3. *Let $\boldsymbol{x} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\boldsymbol{A}$ be a $k \times k$ real symmetric matrix. Then $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \sim \chi_m^2(\lambda)$ with $\lambda = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ if and only if*

$$\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma},\ r(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}) = m,\ \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}.$$

Proof: Let $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{z} + \boldsymbol{\mu}$, where $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}'$, $\boldsymbol{z} \sim N_k(\boldsymbol{0}, \boldsymbol{I})$ and $\tilde{\boldsymbol{A}} = \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}$.

Sufficiency. Note that $\tilde{\boldsymbol{A}}^2 = \tilde{\boldsymbol{A}}$ and $r(\tilde{\boldsymbol{A}}) = m$ due to $\boldsymbol{B} = \boldsymbol{B}\boldsymbol{B}'(\boldsymbol{B}\boldsymbol{B}')^-\boldsymbol{B}$ and

$$r(\tilde{\boldsymbol{A}}) \geq r(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}) = r((\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{B})(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{B})') = r(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{B}) \geq r(\tilde{\boldsymbol{A}}'\tilde{\boldsymbol{A}}) = r(\tilde{\boldsymbol{A}}).$$

Then $||(\boldsymbol{I}-\tilde{\boldsymbol{A}})\boldsymbol{B}'\boldsymbol{A}\boldsymbol{\mu})||^2=\boldsymbol{\mu}'\boldsymbol{A}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma})\boldsymbol{A}\boldsymbol{\mu}=0$, so that $\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}\tilde{\boldsymbol{A}}$ and

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{z}'\tilde{\boldsymbol{A}}\boldsymbol{z}+\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}+2\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{z}=(\boldsymbol{z}+\boldsymbol{c})'\tilde{\boldsymbol{A}}(\boldsymbol{z}+\boldsymbol{c})\sim\chi^2_m(\lambda),$$

where $\boldsymbol{c}'=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}$ and $\lambda=\frac{1}{2}\boldsymbol{c}'\tilde{\boldsymbol{A}}\boldsymbol{c}=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}=\frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$.

Necessity. Suppose $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{z}'\tilde{\boldsymbol{A}}\boldsymbol{z}+\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}+2\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{z}\sim\chi^2_m(\lambda)$. Then

$$(1-2t)^{-\frac{m}{2}}e^{\frac{2t\lambda}{1-2t}}=|\boldsymbol{I}_k-2t\tilde{\boldsymbol{A}}|^{-\frac{1}{2}}e^{t\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}+2t^2\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}(\boldsymbol{I}-2t\tilde{\boldsymbol{A}})^{-1}\boldsymbol{B}'\boldsymbol{A}\boldsymbol{\mu}}$$

for small $|t|$. Comparing the discontinuous points of the two functions on both sides gives

$$(1-2t)^{-\frac{m}{2}}=|\boldsymbol{I}_k-2t\tilde{\boldsymbol{A}}|^{-\frac{1}{2}},$$

which implies that $\tilde{\boldsymbol{A}}^2=\tilde{\boldsymbol{A}}$ and $r(\tilde{\boldsymbol{A}})=m$ (see Section 2), or equivalently

$$\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}=\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\text{ and }r(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma})=m.$$

It follows from above two equations that $\frac{2t\lambda}{1-2t}=t\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}+2t^2\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}(\boldsymbol{I}-2t\tilde{\boldsymbol{A}})^{-1}\boldsymbol{B}'\boldsymbol{A}\boldsymbol{\mu}$, so

$$\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}=2\lambda=(1-2t)[\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}+2t\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}(\boldsymbol{I}-2t\tilde{\boldsymbol{A}})^{-1}\boldsymbol{B}'\boldsymbol{A}\boldsymbol{\mu}],$$

which and $(\boldsymbol{I}-2t\tilde{\boldsymbol{A}})^{-1}=\sum_{n=0}^{\infty}(2t\tilde{\boldsymbol{A}})^n$ imply that for small $|t|$,

$$\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}+2t\boldsymbol{\mu}'(\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}'\boldsymbol{A}-\boldsymbol{A})\boldsymbol{\mu}+4t^2\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}(\tilde{\boldsymbol{A}}-\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}\boldsymbol{\mu}+\cdots$$

By the theory of power series, $\boldsymbol{\mu}'(\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}'\boldsymbol{A}-\boldsymbol{A})\boldsymbol{\mu}=0=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{B}(\tilde{\boldsymbol{A}}-\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}\boldsymbol{\mu}$. That is,

$$\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}=\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}.$$

The proof is completed. □

ACKNOWLEDGMENTS

This research is supported by the Natural Science Foundation of China (NSFC, Grant No. 11561073, 11461079).

REFERENCES

[1] Bonnefond, X. (2012). An analytical proof of Ogawa's determinantal theorem, *Aequationes Mathematicae*, **83**, 127–130.

[2] Carrieu, H. (2010). Close to the Craig-Sakamoto theorem, *Linear Algebra Appl.*, **432**, 777–779.

[3] Carrieu, H. and Lassère, P. (2009). One more simple proof of the Craig-Sakamoto theorem, *Linear Algebra Appl.*, **431**, 1616–1619.

[4] Christensen, R. (2002). *Plane Answers to Complex Questions: The Theory of Linear Models*, 3rd ed., Springer, New York.

[5] Driscoll, M.F. (1999). An improved result relating quadratic forms and chi-square distributions, *American Statistician*, **53**, 273–275.

[6] Driscoll, M.F. and Gundberg Jr., W.R. (1986). A history of the development of Craig's theorem, *American Statistician*, **42**, 139–142.

[7] Driscoll, M.F. and Krasnicka, B. (1995). An accessible proof of Craig's theorem in the general case, *American Statistician*, **49**, 59–62.

[8] Duchesne, P. and Lafaye De Micheaux, P. (2010). Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods, *Computational Statistics and Data Analysis*, **54**, 858–862.

[9] Graybill, F.A. (1976). *Theory and Application of the Linear Model*, North Scituate, MA: Duxbury Press.

[10] Jorgensen, B. (1993). *The Theory of Linear Models*, New York: Chapman & Hall.

[11] Letac, G. and Massam, H. (1995). Craig-Sakamoto's theorem for the Wishart distributions on symmetric cones, *Ann. Inst. Statist. Math.*, **47**, 785–799.

[12] Li, C.K. (2000). A simple proof of the Craig-Sakamoto theorem, *Linear Algebra Appl.*, **321**, 281–283.

[13] Liu, H.; Tang, Y. and Zhang, H.H. (2009). A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, *Computational Statistics and Data Analysis*, **53**, 853–856.

[14] Marcus, M. (1998). On a determinant result of I. Olkin, *Linear Algebra Appl.*, **277**, 237–238.

[15] Mathai, M.A. and Provost, S.B. (1992). *Quadratic Forms in Random Variables, Theory and Applications*, Marcell Dekker, New York.

[16] Matsuura, M. (2003). On the Craig-Sakamoto theorem and Olkin's determinantal result, *Linear Algebra Appl.*, **364**, 321–323.

[17] Ogawa, J. and Olkin, I. (2008). A tale of two countries: The Craig-Sakamoto-Matusita theorem, *J. Statist. Plann. Infer.*, **138**, 3419–3428.

[18] Olkin, I. (1997). A determinantal proof of the Craig-Sakamoto theorem, *Linear Algebra Appl.*, **264**, 217–223.

[19] Provost, S.B. (1996). On Craig's theorem and its generalizations, *J. Statist. Plann. Infer.*, **53**, 311–321.

[20] Ravishanker, N. and Dey, D.K. (2002). *A First Course in Linear Model Theory*, Chapman & Hall/CRC, Boca Raton.

[21] Rao, C.R. (1973). *Linear Statistical Inference and Its Applications* (2nd ed.), New York: Wiley.

[22] Reid, J.G. and Driscoll, M.F. (1988). An accessible proof of Carig's theorem in the non-central case, *American Statistician*, **42**, 139–142.

[23] Searle, S.R. (1972). *Linear Models*, New York: John Wiley & Sons.

[24] Zhang, J. and Yi, J. (2012). A simple proof of the generalized Craig-Sakamoto theorem, *Linear Algebra Appl.*, **437**, 781–782.