International Journal of Statistics and Probability; Vol. 6, No. 6; November 2017
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

# Use of Hotelling's $T^2$: Outlier Diagnostics in Mixtures

D. R. Jensen[1] & D. E. Ramirez[2]

[1] Department of Statistics, Virginia Tech, Blacksburg, VA, USA

[2] Mathematics Department, University of Virginia, Charlottesville, VA, USA

Correspondence: D. E. Ramirez, Mathematics Department, University of Virginia, P. O. Box 400137, Charlottesville, VA 22904, USA. E-mail: der@eservices.virginia.edu

Received: August 9, 2017 Accepted: August 24, 2017 Online Published: September 15, 2017

doi:10.5539/ijsp.v6n6p24 URL: https://doi.org/10.5539/ijsp.v6n6p24

**Abstract**

Given Gaussian observation vectors $[\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]$ having a common mean and dispersion matrix, a pervading issue is to identify shifted observations of type $\{\boldsymbol{Y}_i \to \boldsymbol{Y}_i + \boldsymbol{\delta}_i\}$. Conventional usage enjoins Hotelling's $\mathtt{T}_i^2$ diagnostics, derived and applied under the mutual independence of $[\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]$. Independence often fails, yet the need to identify outliers nonetheless persists. Accordingly, the present study reexamines $\mathtt{T}_i^2$ under dependencies to include equicorrelations and more general matrices. Such dependencies are found in the analysis of calibrated vector measurements and elsewhere. In addition, mixtures of these distributions having star–shaped contours arise on occasion in practice. Nonetheless, the $\mathtt{T}_i^2$ diagnostics are shown to remain exact in level and power for all such mixtures. Moreover, further matrix distributions, not necessarily having finite moments, are seen to generalize $n$–dimensional spherical symmetry to include non–Gaussian matrices of order $(n \times k)$ supporting $\mathtt{T}_i^2$. For these the use of $\mathtt{T}_i^2$ remains exact in level. These findings serve to expand considerably the range of applicability of $\mathtt{T}_i^2$ in practice, to include matrix Cauchy and other heavy tailed distributions intrinsic to econometric and other studies. Case studies serve to illuminate the methodology.

**Keywords** Outlying data, deletion diagnostics, dependent errors, Hotelling's $T^2$

## 1. Introduction

Let $\boldsymbol{Y}_0 = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]'$ comprise $n$ vectors $\{\boldsymbol{Y}_i' = [Y_{i1}, \ldots, Y_{ik}]; 1 \le i \le n\}$ having identical means $\mathtt{E}(\boldsymbol{Y}_i') = \boldsymbol{\mu}' = [\mu_1, \ldots, \mu_k]$ and dispersion matrix $\{\mathtt{V}(\boldsymbol{Y}_i) = \boldsymbol{\Sigma}; 1 \le i \le n\}$. The model is $\{\boldsymbol{Y}_0 = \mathbf{1}_n \cdot \boldsymbol{\mu}' + \mathfrak{E}_0\}$, where $\mathfrak{E}_0 = [\,\epsilon_{ij}]$ consists of random errors, and $\boldsymbol{E}_0 = [\,e_{ij}]$ contains the ordinary observed residuals. The problem at issue, and of persistent concern to users, is whether shifts of type $\{\boldsymbol{Y}_i \to \boldsymbol{Y}_i + \boldsymbol{\delta}_i\}$ in $\mathbb{R}^k$ may have occurred. Numerous approaches have been advocated, to include graphics and numerical diagnostics. Selected references are listed subsequently; a recent survey is Rodrigues and Boente (2011). Prominent deletion diagnostics are modeled on Mahalanobis (1936) distance metrics in $\mathbb{R}^k$.

Under single–case deletions, the rows $[\boldsymbol{Y}_i', \boldsymbol{\epsilon}_i']$ are removed from $[\boldsymbol{Y}_0, \mathfrak{E}_0]$, retaining $\boldsymbol{Y}$ from $\boldsymbol{Y}_0 = [\boldsymbol{Y}', \boldsymbol{Y}_i]'$ and $\boldsymbol{e}_i$ from $\boldsymbol{E}_0 = [\boldsymbol{E}', \boldsymbol{e}_i]'$. For $k = 1$, with $S_i^2$ as the residual mean square, the $R$–Student statistics $t_i^2 = n\, e_i^2/(n-1)S_i^2$ trace to Snedecor and Cochran (1968, page 157) in testing for a single shift $\{Y_i + \delta_i\}$; see also Beckman and Trussell (1974). Corresponding to $t_i^2$ for $k > 1$ are Hotelling's (1931) diagnostics $\mathtt{T}_i^2$ given by

$$\left\{ \mathtt{T}_i^2 = \frac{n\, \boldsymbol{e}_i' \boldsymbol{S}_i^{-1} \boldsymbol{e}_i}{(n-1)};\ 1 \le i \le n \right\} \tag{1}$$

where $(n-1)\boldsymbol{S}_i = \boldsymbol{Y}'\boldsymbol{Y}$. In having exact and well documented normal–theory operating characteristics, these remain the diagnostics of choice for many users. Initially derived under normality and the mutual independence of $[\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]$, these assumptions continue to validate the use of $\mathtt{T}_i^2$ up to the present. For early developments see Caroni (1987), and more recently Barrett and Ling (1992) and Barrett (2003), together with references cited. Clearly independence often fails in practice, yet the need to identify outliers nonetheless persists.

As background, precedents for this study in the case $k = 1$ include $t_i^2$ and the diagnostics of Dixon (1950), Grubbs (1950), and Ferguson (1961) based on order statistics. As reassessed in Jensen and Ramirez (2015), all remain exact in level and power under dependent ensembles of distributions in $\mathbb{R}^n$, and for mixtures over these ensembles having star–shaped contours.

The *objectives* of this research are to extend those findings to include mixture distributions for Gaussian matrices $\boldsymbol{Y}_0 = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]'$ having dependent elements with the direct–product structure $\mathtt{V}(\boldsymbol{Y}_0) = \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}$, and for mixtures of these. Here $\boldsymbol{\Omega}$ is taken as an equicorrelation matrix $\boldsymbol{\Omega}(\rho)$ or the more general $\boldsymbol{\Omega}(\boldsymbol{\xi})$ to be identified. In addition, left–spherical matrix distributions serve to generalize the notion of spherical symmetry on $\mathbb{R}^n$, thereby to encompass matrix stable laws and Cauchy distributions. For the latter the use of $\mathtt{T}_i^2$ nonetheless is seen to remain exact in level. These findings serve to expand considerably the range of applicability of $\mathtt{T}_i^2$ in practice, to include heavy tailed distributions. Accordingly, normal-theory $\mathtt{T}_i^2$ diagnostics are seen to be genuinely nonparametric.

Precedents for undertaking the mixtures and dependence structures of this study trace to Box and Tiao (1968) and Aitken and Wilson (1980), who modeled data from subsamples as Gaussian mixtures. Moreover, among other venues, calibrated data subject to errors of calibration often are equicorrelated under both direct and inverse calibration, as seen in Jensen and Ramirez (2009, 2012). The importance of heavy–tailed distributions in economics and finance is highlighted in Ibragimov *et al.* (2015). An outline of the study follows.

Preliminary developments are given in Section 2. The principal findings follow in Section 3, and some consequences of these are detailed through examples in Section 4. Critical supporting topics, to include essential matrix distributions, are attached for completeness as an Appendix.

## 2. Preliminaries

### 2.1 Notation

Spaces of note include the Euclidean $n$-space $\mathbb{R}^n$; its positive orthant $\mathbb{R}^n_+$; the real $(n\times k)$ matrices $\mathbb{F}_{n\times k}$; the symmetric $(n\times n)$ matrices $\mathbb{S}_n$; and their positive definite varieties $\mathbb{S}_n^+$. Vectors and matrices are set in bold type; the transpose, inverse, trace, and determinant of $\boldsymbol{A}$ are $\boldsymbol{A}'$, $\boldsymbol{A}^{-1}$, tr$(\boldsymbol{A})$, and $|\boldsymbol{A}|$; $\boldsymbol{I}_n$ is the $(n\times n)$ identity; Diag$(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k)$ is a block-diagonal array; and $\mathbf{1}_n' = [1,\ldots,1] \in \mathbb{R}^n$ is the unit vector. The direct matrix product is $\boldsymbol{A}\otimes\boldsymbol{B} = [a_{ij}\boldsymbol{B}]$. An idempotent matrix of note is $\boldsymbol{B}_n = (\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')$.

A random $\boldsymbol{Y}\in\mathbb{R}^n$ has distribution $\mathcal{L}(\boldsymbol{Y})$; the mean vector $\mathtt{E}(\boldsymbol{Y})$ and dispersion matrix $\mathtt{V}(\boldsymbol{Y}) = \boldsymbol{\Sigma}$, with variance Var$(Y) = \sigma^2$ on $\mathbb{R}^1$. Its density $(pdf)$ and cumulative distribution function $(cdf)$ are $g(\boldsymbol{y})$ and $G(\boldsymbol{y})$; and its characteristic function $(chf)$ is $\phi_{\boldsymbol{Y}}(\boldsymbol{t})$. Specifically, $\mathcal{L}(\boldsymbol{Y}) = N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is Gaussian in $\mathbb{R}^n$ with mean vector $\boldsymbol{\mu}$ and dispersion matrix $\boldsymbol{\Sigma}$, whereas $\mathcal{L}(\boldsymbol{Y}_0) = N_{n\times k}(\boldsymbol{M}, \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$ is Gaussian in $\mathbb{F}_{n\times k}$ with designated parameters. A random $\boldsymbol{W}\in\mathbb{S}_k^+$ is said to have the Wishart distribution $\mathtt{W}_k(\nu, \boldsymbol{\Sigma}, \boldsymbol{\Theta})$ of order $k$, with $\nu$ degrees of freedom, the scale parameters $\boldsymbol{\Sigma}$, and the noncentrality matrix $\boldsymbol{\Theta}$. Further details are supplied in Appendix A.1.

Distributions on $\mathbb{R}^1_+$ include $\chi^2(u;\nu,\sigma^2,\lambda)$ as chi-squared with argument $u$, having $\nu$ degrees of freedom, the scale parameter $\sigma^2$, and noncentrality $\lambda$; and Hotelling's (1931) $T_k^2(u;\nu,\lambda)$ of order $k$ having $\nu$ degrees of freedom and noncentrality parameter $\lambda$. Moreover, $F(u;\nu_1,\nu_2,\lambda)$ is the noncentral Snedecor–Fisher $F$–distribution having $\nu_1$ and $\nu_2$ degrees of freedom, which increases stochastically with the noncentrality parameter $\lambda$. Identify $\{\mathtt{T}_i^2 > c_\alpha\}$ as the conventional $\alpha$–level rejection rule based on $T_k^2(u;\nu,0)$.

### 2.2 The Model

We specialize from the model $\{\boldsymbol{Y}_0 = \boldsymbol{X}_0\boldsymbol{\mathcal{B}} + \mathfrak{E}_0\}$ with $\boldsymbol{Y}_0 = [\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_n]' \in \mathbb{F}_{n\times k}$, $\boldsymbol{X}_0 = [\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n]' \in \mathbb{F}_{n\times d}$ of rank $d < n$ with $\boldsymbol{x}_i$ as design point $i$, and $\boldsymbol{\mathcal{B}} = [\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_k] \in \mathbb{F}_{d\times k}$. Under single–case deletions $\mathfrak{E}_0 = [\mathfrak{E}', \boldsymbol{\epsilon}_i]'$ consists of random errors; $\boldsymbol{E}_0 = [\boldsymbol{E}', \boldsymbol{e}_i]'$ are ordinary residuals; and $\boldsymbol{H}_n = \boldsymbol{X}_0(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{X}_0'$.

**Definition 1.** In particular, we take $\{\boldsymbol{Y}_0 = \mathbf{1}_n\cdot\boldsymbol{\mu}' + \mathfrak{E}_0\}$ and $\boldsymbol{H}_n = \mathbf{1}_n\mathbf{1}_n'/n$ in keeping with the objectives of this study.

To continue, extended Gauss-Markov assumptions take $\mathtt{V}(\boldsymbol{Y}_0) = \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}$, where $\boldsymbol{\Omega}$ takes values $\boldsymbol{\Omega}(\rho)$ (equivalently $\boldsymbol{\Omega}(\theta)$) and $\boldsymbol{\Omega}(\boldsymbol{\xi})$ as in Section 2.3.

**Assumptions A.** The following hold for a model with a single shift.

$A_1$. $\mathtt{E}(\mathfrak{E}_0) = \boldsymbol{\Delta} = [\boldsymbol{O}', \boldsymbol{\delta}_i]'$ with $\boldsymbol{O}(n-1\times k)$ a matrix of zeroes such that $\mathtt{E}(\mathfrak{E}) = \boldsymbol{O}$ and $\mathtt{E}(\boldsymbol{\epsilon}_i) = \boldsymbol{\delta}_i$;

$A_2$. $\mathtt{V}(\mathfrak{E}_0) = \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}$ with $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\rho), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$; and

$A_3$. $\mathcal{L}(\mathfrak{E}_0) = N_{n\times k}(\boldsymbol{\Delta}, \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$ for $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\rho), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$.

As in conventional deletion diagnostics, this represents a shift $\{\boldsymbol{Y}_i \to \boldsymbol{Y}_i + \boldsymbol{\delta}_i\}$ at the design point $\boldsymbol{x}_i$ (now the index $i$) in $\boldsymbol{X}_0$.

### 2.3 The Matrices $\boldsymbol{\Xi}$

Validity in linear inference rests in part on the structure of dispersion matrices. Three cases are considered, namely $\boldsymbol{\Omega}(\theta)$, $\boldsymbol{\Omega}(\boldsymbol{\xi})$, and $\boldsymbol{\Omega}(\rho)$ where, for $\boldsymbol{\xi}' = [\xi_1,\ldots,\xi_n]$, we have $\tau_1 = \xi_1 + \ldots + \xi_n = n\bar{\xi}$ and $\tau_2 = \sum_{i=1}^n(\xi_i - \bar{\xi})^2$. Details follow, where

$$\alpha_1 = \frac{1}{2}[\tau_1 + (\tau_1^2 + 4n\tau_2)^{\frac{1}{2}}] \quad \text{and} \quad \alpha_n = \frac{1}{2}[\tau_1 - (\tau_1^2 + 4n\tau_2)^{\frac{1}{2}}]. \tag{2}$$

**Lemma 1.** *(i) Let* $\mathbf{\Omega}(\theta) = \sigma^2(\boldsymbol{I}_n + \theta\mathbf{1}_n\mathbf{1}_n')$*; its eigenvalues are 1.0, with multiplicity* $n-1$*, and* $1+n\theta$*, so that* $\mathbf{\Omega}(\theta)$ *is positive definite if and only if* $\theta \in \Gamma_1 = \{\theta : \theta > -\frac{1}{n}\}$.

*(ii) Let* $\mathbf{\Omega}(\boldsymbol{\xi}) = \sigma^2(\boldsymbol{I}_n + \mathbf{1}_n\boldsymbol{\xi}' + \boldsymbol{\xi}\mathbf{1}_n' - \bar{\xi}\mathbf{1}_n\mathbf{1}_n')$ *with* $\mathbf{0} \neq \boldsymbol{\xi} \neq \theta\mathbf{1}_n$*; its ordered eigenvalues are* $\{\kappa_1 = 1+\alpha_1, \kappa_2 = \ldots = \kappa_{n-1} = 1, \kappa_n = 1+\alpha_n\}$ *as in (2.1); then* $\mathbf{\Omega}(\boldsymbol{\xi})$ *is positive definite if and only if* $\boldsymbol{\xi} \in \Gamma_2 = \{\boldsymbol{\xi} \in \mathbb{R}^n : \tau_1 > n\tau_2 - 1\}$.

*(iii) Let* $\mathbf{\Omega}(\rho) = \sigma^2[(1-\rho)\boldsymbol{I}_n + \rho\mathbf{1}_n\mathbf{1}_n']$*, the equicorrelated case; then* $\mathbf{\Omega}(\rho)$ *is positive definite if and only if* $\rho \in \Gamma_3 = \{\rho : -\frac{1}{n-1} < \rho < 1\}$.

**Proof.** Details are given in Jensen (1996).■

That $\mathbf{\Omega}(\theta)$ and $\mathbf{\Omega}(\rho)$ are equivalent follows on taking $\theta = \rho/(1-\rho)$, so that $\mathbf{\Omega}(\theta) = \frac{1}{1-\rho}\mathbf{\Omega}(\rho)$. Accordingly, the collections $\Xi_1 = \{\mathbf{\Omega}(\theta); \theta \in \Gamma_1\}$, equivalently, $\Xi_1 = \{\mathbf{\Omega}(\rho); \rho \in \Gamma_3\}$, and $\Xi_2 = \{\mathbf{\Omega}(\boldsymbol{\xi}); \boldsymbol{\xi} \in \Gamma_2\}$, comprise ensembles of positive definite matrices, to be amalgamated as $\Xi = \Xi_1 \cup \Xi_2$. For further details see Jensen (1996).

*2.4 Mixture Distributions*

From Assumption $A_1$ let $\mathbf{\Lambda} = \mathbf{1}_n \cdot \boldsymbol{\mu}' + \mathbf{\Delta}$ and consider $g_{n\times k}(\boldsymbol{y}; \mathbf{\Lambda}, \mathbf{\Omega} \otimes \mathbf{\Sigma})$ in $\mathbb{F}_{n\times k}$ as the Gaussian density corresponding to $N_{n\times k}(\mathbf{\Lambda}, \mathbf{\Omega} \otimes \mathbf{\Sigma})$ as in Appendix A.1. These generate ensembles as $\mathbf{\Omega}$ ranges over $\Xi$, namely

$$E_1 = \{g_{n\times k}(\boldsymbol{y}; \mathbf{\Lambda}, \mathbf{\Omega}(\theta) \otimes \mathbf{\Sigma}); \theta \in \Gamma_1\} \tag{3}$$

$$E_2 = \{g_{n\times k}(\boldsymbol{y}; \mathbf{\Lambda}, \mathbf{\Omega}(\boldsymbol{\xi}) \otimes \mathbf{\Sigma}); \boldsymbol{\xi} \in \Gamma_2\}. \tag{4}$$

Next visualize the ensemble $E_1$ to have mixing parameters $\theta$, and $E_2$ to have mixing parameters $\boldsymbol{\xi}$. Then mixtures in $\mathbb{F}_{n\times k}$ of type

$$f_i(\boldsymbol{y}; \mathbf{\Lambda}, G_i) = \int_{\Gamma_i} g_{n\times k}(\boldsymbol{y}; \mathbf{\Lambda}, \mathbf{\Omega}(i) \otimes \mathbf{\Sigma}) dG_i(\cdot) \tag{5}$$

emerge with $G_i \in \{G_1, G_2\}$ as $cdf$s on $\Gamma_i \in \{\Gamma_1, \Gamma_2\}$, and with $\mathbf{\Omega}(i) \in [\mathbf{\Omega}(\theta), \mathbf{\Omega}(\boldsymbol{\xi})]$. In particular, the densities $f_1(\boldsymbol{y}; \mathbf{\Lambda}, G_1)$ and $f_2(\boldsymbol{y}; \mathbf{\Lambda}, G_2)$ are dispersion mixtures of elliptical Gaussian distributions on $\mathbb{F}_{n\times k}$ centered at $\mathbf{\Lambda} \in \mathbb{F}_{n\times k}$. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ comprise all $cdf$s on $\Gamma_1$ and $\Gamma_2$, respectively; these in turn generate the collections

$$\mathrm{M}_1 = \{f_1(\boldsymbol{y}; \mathbf{\Lambda}, G_1);\ G_1 \in \mathcal{G}_1\} \tag{6}$$

$$\mathrm{M}_2 = \{f_2(\boldsymbol{y}; \mathbf{\Lambda}, G_2);\ G_2 \in \mathcal{G}_2\} \tag{7}$$

comprising all dispersion mixtures of the referenced types.

## 3. The Principal Findings

*3.1 Overview*

Taking $\{\boldsymbol{Y}_0 = \mathbf{1}_n \cdot \boldsymbol{\mu}' + \mathfrak{E}_0\}$ in $\mathbb{F}_{n\times k}$ as in Definition 1, we rearrange elements such that $\boldsymbol{Y}_0 = [\boldsymbol{Y}', \boldsymbol{Y}_i]'$, $\mathfrak{E}_0 = [\mathfrak{E}', \boldsymbol{\epsilon}_i]'$, and $\boldsymbol{E}_0 = [\boldsymbol{E}', \boldsymbol{e}_i]'$, and consider arbitrary shifts $\{\boldsymbol{Y}_i \to \boldsymbol{Y}_i + \boldsymbol{\delta}_i\}$. Nonstandard versions of $\mathcal{L}(\mathrm{T}_i^2)$ as in expression (1) are to be studied, but where the independence of $\{\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n\}$ fails. Instead take $\mathcal{L}(\boldsymbol{Y}_0) = N_{n\times k}(\boldsymbol{M}, \mathbf{\Omega} \otimes \mathbf{\Sigma})$ with $\mathbf{\Omega} \in \{\mathbf{\Omega}(\rho), \mathbf{\Omega}(\boldsymbol{\xi})\}$, equivalently $\mathbf{\Omega} \in \{\mathbf{\Omega}(\theta), \mathbf{\Omega}(\boldsymbol{\xi})\}$, as the basic model undergirding the $\mathrm{T}_i^2$ diagnostics. Here $\boldsymbol{M} = \mathbf{\Lambda} = \mathbf{1}_n \cdot \boldsymbol{\mu}' + \mathbf{\Delta}$ from Assumption $A_1$. It is seen that shifts $\{\boldsymbol{Y}_i \to \boldsymbol{Y}_i + \boldsymbol{\delta}_i\}$ propagate into noncentralitiy parameters of $\mathcal{L}(\mathrm{T}_i^2)$. These findings in turn rest on matrices of quadratic and bilinear forms of type $\boldsymbol{Y}_0'\boldsymbol{A}\boldsymbol{Y}_0$ in the observed data matrix $\boldsymbol{Y}_0 \in \mathbb{F}_{n\times k}$ as in Appendix A.1. Here $\boldsymbol{Y}_i$ is a generic test case; another would entail rearranging $\boldsymbol{Y}_0$ as $\boldsymbol{Y}_0 = [\boldsymbol{Y}', \boldsymbol{Y}_j]'$.

*3.2 Properties of Residuals*

The observed residuals $\boldsymbol{E}_0$ under Assumptions A are germane, as $\mathrm{T}_i^2$ is a function of these. In particular, it remains to evaluate $\mathrm{E}(\boldsymbol{e}_i)$, $\mathrm{Var}(\boldsymbol{e}_i)$ and $\mathcal{L}(\boldsymbol{e}_i)$ as special cases. Details follow, where $r = (n-1)$ and $\mathrm{E}(\mathfrak{E}_0) = \mathbf{\Delta} = [\mathbf{0}', \boldsymbol{\delta}_i]'$ as Assumption $A_1$, and $\boldsymbol{B}_n = (\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')$.

**Theorem 1.** *Consider the ordinary residuals* $\boldsymbol{E}_0 = [\boldsymbol{E}', \boldsymbol{e}_i]'$ *under* Assumptions A, *with* $\mathbf{\Omega} \in \{\mathbf{\Omega}(\theta), \mathbf{\Omega}(\boldsymbol{\xi})\}$*, and let* $T(\boldsymbol{E}_0)$ *be a mapping to a linear space* $\mathbb{V}$*. Then the following properties hold independently of* $\mathbf{\Omega} \in \{\mathbf{\Omega}(\theta), \mathbf{\Omega}(\boldsymbol{\xi})\}$.

*(i)* $\mathrm{V}(\boldsymbol{E}_0) = \boldsymbol{B}_n \otimes \mathbf{\Sigma}$ *and* $\mathcal{L}(\boldsymbol{E}_0) = N_{n\times k}(\mathbf{\Delta}, \boldsymbol{B}_n \otimes \mathbf{\Sigma})$;

*(ii)* $\mathrm{E}(\boldsymbol{e}_i) = \frac{r}{n}\boldsymbol{\delta}_i$, $\mathrm{V}(\boldsymbol{e}_i) = \frac{r}{n}\mathbf{\Sigma}$, *and* $\mathcal{L}(\boldsymbol{e}_i) = N_k(\frac{r}{n}\boldsymbol{\delta}_i, \frac{r}{n}\mathbf{\Sigma})$;

*(iii)* $\mathcal{L}(T(\boldsymbol{E}_0|\mathbf{\Omega})) = \mathcal{L}(T(\boldsymbol{E}_0 | \boldsymbol{I}_n))$.

**Proof.** Observe from $\boldsymbol{E}_0 = \boldsymbol{B}_n\boldsymbol{Y}_0$ and the conventions of Appendix A.1 that $\mathtt{V}(\boldsymbol{E}_0) = \boldsymbol{B}_n\boldsymbol{\Omega}\boldsymbol{B}_n\otimes\boldsymbol{\Sigma} = \boldsymbol{B}_n\otimes\boldsymbol{\Sigma}$ for $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\theta), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$, since $\boldsymbol{B}_n\boldsymbol{1}_n = \boldsymbol{0}$ annihilates successive terms in expansions for $\boldsymbol{\Omega}(\theta)$ and $\boldsymbol{\Omega}(\boldsymbol{\xi})$ in the product terms following the first. This together with Assumption $A_3$ gives (i). The expected product $\mathtt{E}(\boldsymbol{E}_0) = (\boldsymbol{I}_n - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n')\,\boldsymbol{\Delta}$ in partitioned form is

$$\mathtt{E}\begin{bmatrix}\boldsymbol{E}\\ \boldsymbol{e}_i'\end{bmatrix} = \frac{1}{n}\begin{bmatrix}(n\boldsymbol{I}_r - \boldsymbol{1}_r\boldsymbol{1}_r') & -\boldsymbol{1}_r\\ -\boldsymbol{1}_r' & r\end{bmatrix}\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{\delta}_i'\end{bmatrix} = \frac{1}{n}\begin{bmatrix}-\boldsymbol{1}_r\boldsymbol{\delta}_i'\\ r\boldsymbol{\delta}_i'\end{bmatrix}. \tag{8}$$

Thus $\mathtt{E}(\boldsymbol{e}_i) = r\boldsymbol{\delta}_i/n$ and $\mathtt{V}(\boldsymbol{e}_i) = \frac{r}{n}\boldsymbol{\Sigma}$ is the $(n, n)$ block of $\boldsymbol{B}_n \otimes \boldsymbol{\Sigma}$ which, together with normality, give conclusion (ii). Conclusion (iii) follows directly. ■

*3.3 Nonstandard Matrix Forms*

Generalizing from Lemma A.1(iii) of Jensen (2001a) and from Jensen and Ramirez (2014) in extending to include matrix arrays, the multivariate Fisher–Cochran expansion generating $\mathtt{T}_i^2$ is

$$\boldsymbol{Y}_0'\boldsymbol{A}_1\boldsymbol{Y}_0 + \boldsymbol{Y}_0'\boldsymbol{A}_2\boldsymbol{Y}_0 = \boldsymbol{Y}_0'\boldsymbol{A}_3\boldsymbol{Y}_0; \tag{9}$$

$$\frac{n}{r}\boldsymbol{e}_i\boldsymbol{e}_i' + (r-1)\boldsymbol{S}_i = \boldsymbol{E}_0'\boldsymbol{E}_0; \tag{10}$$

where $(\boldsymbol{e}_i, \boldsymbol{\delta}_i)$ are of order $(k\times 1)$, and the second line explains the first. Moreover, $(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3)$ are idempotent as given explicitly in Jensen and Ramirez (2015), namely, $\boldsymbol{A}_3 = \boldsymbol{B}_n = (\boldsymbol{I}_n - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n')$, $\boldsymbol{A}_2 = \mathrm{Diag}(\boldsymbol{B}_r, \boldsymbol{0})$ with $\boldsymbol{B}_r = (\boldsymbol{I}_r - \frac{1}{r}\boldsymbol{1}_r\boldsymbol{1}_r')$, and $\boldsymbol{A}_1 = \boldsymbol{A}_3 - \boldsymbol{A}_2$. To continue, designate the aforementioned matrix forms as $\boldsymbol{Q}_1 = \boldsymbol{Y}_0'\boldsymbol{A}_1\boldsymbol{Y}_0$, $\boldsymbol{Q}_2 = \boldsymbol{Y}_0'\boldsymbol{A}_2\boldsymbol{Y}_0$, and $\boldsymbol{Q}_3 = \boldsymbol{Y}_0'\boldsymbol{A}_3\boldsymbol{Y}_0$, and recall that $\boldsymbol{M} = \boldsymbol{\Lambda} = \boldsymbol{1}_n\cdot\boldsymbol{\mu}' + \boldsymbol{\Delta}$.

**Theorem 2.** *Given $\mathcal{L}(\boldsymbol{Y}_0) = N_{n\times k}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$ under* Assumptions A*; take $\boldsymbol{E}_0 = \boldsymbol{B}_n\boldsymbol{Y}_0$; and consider the matrix forms $\{\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_3\}$ as in (9). Then despite dependencies among elements of $\boldsymbol{Y}_0 = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]'$ we have for each $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\theta), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$ the following.*

*(i) $\mathcal{L}(\boldsymbol{Q}_1) = \mathbb{W}_k(1; \boldsymbol{\Sigma}, \boldsymbol{\Theta}_i)$, $\boldsymbol{\Theta}_i = \frac{r}{n}\boldsymbol{\delta}_i\boldsymbol{\delta}_i'$;*

*(ii) $\mathcal{L}(\boldsymbol{Q}_2) = \mathbb{W}_k(r-1; \boldsymbol{\Sigma}, \boldsymbol{0})$;*

*(iii) $\mathcal{L}(\boldsymbol{Q}_3) = \mathbb{W}_k(r; \boldsymbol{\Sigma}, \boldsymbol{\Theta}_i)$;*

*(iv) $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are distributed independently;*

*(v) $\mathtt{T}_i^2 = (r-1)\,\mathrm{tr}\,\boldsymbol{Q}_1\boldsymbol{Q}_2^{-1}$;*

*(vi) $\{\mathcal{L}(\mathtt{T}_i^2) = T_k^2(u; r-1, \lambda_i)$; $\lambda_i = \frac{r}{n}\boldsymbol{\delta}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_i$; $1\le i\le n\}$.*

**Proof.** Fix $\boldsymbol{u}\in\mathbb{R}^k$ and let $\{Q_i^{\boldsymbol{u}} = \boldsymbol{u}'\boldsymbol{Y}_0'\boldsymbol{A}_i\boldsymbol{Y}_0\boldsymbol{u}; i = 1, 2, 3\}$, such that $\mathcal{L}(\boldsymbol{u}'\boldsymbol{Y}_0) = N_n(\boldsymbol{u}'\boldsymbol{\Lambda}, \sigma_{\boldsymbol{u}}^2\boldsymbol{\Omega})$ with $\sigma_{\boldsymbol{u}}^2 = \boldsymbol{u}'\boldsymbol{\Sigma}\boldsymbol{u}$. Drawing on Mathai and Provost (1992, page 201), Jensen and Ramirez (2015) established conclusions (i)–(iv) for $(Q_1^{\boldsymbol{u}}, Q_2^{\boldsymbol{u}}, Q_3^{\boldsymbol{u}}\}$ in terms of the corresponding $\chi^2$ distributions on demonstrating that $\{\boldsymbol{A}_i\boldsymbol{\Omega}\boldsymbol{A}_i = \boldsymbol{A}_i; i = 1, 2, 3\}$ and that $\boldsymbol{A}_1\boldsymbol{\Omega}\boldsymbol{A}_2 = \boldsymbol{0}$ for $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\theta), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$, as in Appendix A.1 of Jensen and Ramirez (2015). Theorem A.1 of the attached Appendix now lifts those results to encompass the Wishart distributions of conclusions (i)–(iv). Conclusion (v) follows directly from expressions (1) and (10), and conclusion (vi) as the Wishart analog of the noncentral properties of Theorem A.1 of Jensen and Ramirez (2015). ■

*3.4 Invariance under Mixtures*

That the $\mathtt{T}_i^2$ diagnostics may be valid under star–contoured errors is the subject of the following, where $\boldsymbol{\Lambda} = \boldsymbol{1}_n\cdot\boldsymbol{\mu}' + \boldsymbol{\Delta}$ as in Assumption $A_1$.

**Theorem 3.** *Given the model $\{\boldsymbol{Y}_0 = \boldsymbol{\Lambda} + \mathfrak{E}_0\}$ having a Gaussian mixture density $f_1(\boldsymbol{y}_0; \boldsymbol{\Lambda}, G_1)$ in $M_1$ or $f_2(\boldsymbol{y}_0; \boldsymbol{\Lambda}, G_2)$ in $M_2$ as in Section 2.4.*

*(i) Tests using $\mathtt{T}_i^2$ remain exact in level and power for all mixtures in $M_1$;*

*(ii) Tests using $\mathtt{T}_i^2$ remain exact in level and power for all mixtures in $M_2$;*

*(iii) These $\mathtt{T}_i^2$ distributions are identical to those derived from $\mathcal{L}(\boldsymbol{Y}_0) = N_{n\times k}(\boldsymbol{\Lambda}, \boldsymbol{I}_n\otimes\boldsymbol{\Sigma})$.*

**Proof.** Return to Section 2.4 and expression (5). We argue conditionally as follows: (i) Fix $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\theta), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$; (ii) note from Theorem 3.2(vi) that $\mathcal{L}(\mathtt{T}_i^2)$ holds independently of $\boldsymbol{\Omega}$; then (iii) make the change–of–variables behind the integral, to conclude for $\{i=1,2\}$ that

$$f_i(\boldsymbol{y}_0, \boldsymbol{\Lambda}, G_i) = \int_{\Gamma_i} g_{n\times k}(\boldsymbol{y}_0; \boldsymbol{\Lambda}, \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})dG_i(\cdot)$$
$$f_i(\boldsymbol{y}_0, \boldsymbol{\Lambda}, G_i) \rightarrow \int_{\Gamma_i} T_k^2(u; r-1, \lambda_i)dG_i(\cdot) = T_k^2(u; r-1, \lambda_i)$$

independently of $\boldsymbol{\Omega}\in\{\boldsymbol{\Omega}(\theta), \boldsymbol{\Omega}(\boldsymbol{\xi})\}$ and of $G_i$, and since $\int_{\Gamma_i} dG_i = 1$. ■

*3.5 Left–Spherical Matrix Distributions*

Other structured distributions are germane. This section draws heavily from Jensen and Good (1981).

**Definition 2.** A distribution on $\mathbb{F}_{n\times k}$ is *left–spherical* provided that $\mathcal{L}(\boldsymbol{X}) = \mathcal{L}(\boldsymbol{PX})$ for every orthogonal matrix $\boldsymbol{P}\in O(n)$. Denote by $L_{n\times k}(\boldsymbol{0}, \boldsymbol{I}_n)$ the class of left–spherical matrix distributions centered at the origin in $\mathbb{F}_{n\times k}$, in which case its *chf* has the form $\phi_{\boldsymbol{X}} = \psi(\mathrm{tr}\, \boldsymbol{T}'\boldsymbol{T})$ with argument $\boldsymbol{T}\in\mathbb{F}_{n\times k}$.

Nor are these required to have moments of various orders: examples are the *left–spherical stable laws* on $\mathbb{F}_{n\times k}$ having *chf*s of type

$$\phi_{\boldsymbol{X}}(\boldsymbol{T}) = \exp(\gamma[\mathrm{tr}(\boldsymbol{T}'\boldsymbol{T})]^{\frac{\alpha}{2}}) \tag{11}$$

with parameters $\{\gamma<0,\ 0<\alpha<2\}$ of which the matrix Cauchy law with $\alpha=1$ is a noteworthy special case. In addition, the shift $\{\boldsymbol{X}\rightarrow\boldsymbol{Z}=\boldsymbol{X}+\mathtt{M}\}$ gives $\mathcal{L}(\boldsymbol{Z}) = L_{n\times k}(\mathtt{M}, \boldsymbol{I}_n)$.

To continue, let $\mathfrak{M}$ be a linear subspace of $\mathbb{F}_{n\times k}$. The following is fundamental.

**Definition 3.** The transformation $T : \mathbb{F}_{n\times k}\rightarrow\mathcal{V}$ is said to be *translation–invariant* with respect to $\mathfrak{M}$ if, for each $\boldsymbol{Z}\in\mathbb{F}_{n\times k}$ and $\mathtt{M}\in\mathfrak{M}$, $T(\boldsymbol{Z}+\mathtt{M}) = T(\boldsymbol{Z})$. In addition, $T$ is *right–invariant* under $\mathcal{G}l(k)$ if, for every $\boldsymbol{B}\in\mathcal{G}l(k)$, $T(\boldsymbol{ZB}) = T(\boldsymbol{Z}.)$

The following is given as Theorem 2 of Jensen and Good (1981).

**Lemma 2.** *Suppose $\mathcal{L}(\boldsymbol{Z})\in\{L_{n\times k}(\mathtt{M}, \boldsymbol{I}_n); \mathtt{M}\in\mathfrak{M}\}$, and let the transformation $T : \mathbb{F}_{n\times k}\rightarrow\mathcal{V}$ be translation–invariant with respect to $\mathfrak{M}$ and be right–invariant under $\mathcal{G}l(k)$. Then the distribution of $T(\boldsymbol{Z})$ is the same for all $\mathcal{L}(\boldsymbol{Z})\in L_{n\times k}(\mathtt{M}, \boldsymbol{I}_n)$ independently of $\mathtt{M}$, and thus is identical to its matrix normal theory form.*

We next examine $\mathtt{T}_i^2$ under $\mathcal{L}(Y_0) = L_{n\times k}(\boldsymbol{\Lambda}, \boldsymbol{I}_n)$, where $\boldsymbol{\Lambda} = \boldsymbol{1}_n\cdot\boldsymbol{\mu}' + [\boldsymbol{O}',\ \boldsymbol{\delta}_i]'$ with $\boldsymbol{O}(r\times k)$ a matrix of zeroes as in Assumption $A_1$. We proceed in two steps, first taking $\boldsymbol{Y}_0\rightarrow\boldsymbol{GY}_0$, then the latter into $\mathtt{T}_i^2$, namely,

$$\boldsymbol{GY}_0 = \begin{bmatrix} \boldsymbol{B}_r & \boldsymbol{0} \\ -\frac{1}{n}\boldsymbol{1}_r' & (1-\frac{1}{n}) \end{bmatrix}\begin{bmatrix} \boldsymbol{Y} \\ \boldsymbol{Y}_i' \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_r\boldsymbol{Y} \\ \boldsymbol{e}_i' \end{bmatrix} \rightarrow \frac{n}{r}\,\boldsymbol{e}_i'[\boldsymbol{Y}'\boldsymbol{B}_r\boldsymbol{Y}]^{-1}\boldsymbol{e}_i = \mathtt{T}_i^2. \tag{12}$$

As $\boldsymbol{B}_r\boldsymbol{Y}$ contains deviations from means, it is clear that $(n-1)\boldsymbol{S}_i = \boldsymbol{Y}'\boldsymbol{B}_r\boldsymbol{Y}$. In addition, the final row of $\boldsymbol{G}$ is $\boldsymbol{u}' = [-\frac{1}{n},\ldots,-\frac{1}{n},(1-\frac{1}{n})]$, so that $\boldsymbol{u}'\boldsymbol{Y}_0 = \boldsymbol{e}_i'$. Moreover, under Assumptions A, $\mathrm{E}(\boldsymbol{B}_r\boldsymbol{Y}) = \boldsymbol{O}$ and $\mathrm{E}(\boldsymbol{e}_i) = \boldsymbol{\delta}_i$.

To continue, let $\mathfrak{M} = \{\boldsymbol{1}_n\cdot\boldsymbol{\mu}'\ |\boldsymbol{\mu}\in\mathbb{R}^k\}$. We have the following.

**Theorem 4.** *Suppose that $\mathcal{L}(Y_0)\in\{L_{n\times k}(\mathtt{M}, \boldsymbol{I}_n); \mathtt{M}\in\mathfrak{M}\}$, and consider testing $H_0:\boldsymbol{\delta}_i=\boldsymbol{0}$ against $H_1:\boldsymbol{\delta}_i\neq\boldsymbol{0}$.*

*(i) Then the null distribution of $\mathtt{T}_i^2$ is the same for all $\mathcal{L}(Y_0)\in\{L_{n\times k}(\mathtt{M}, \boldsymbol{I}_n); \mathtt{M}\in\mathfrak{M}\}$ and thus is identical to its matrix normal theory form.*

*(ii) This holds in particular for every left–spherical stable law on $\mathbb{F}_{n\times k}$ having moments of order up to but excluding $\alpha$.*

**Proof.** Conclusion (i) follows directly from Lemma 2, where the translation invariance of $\mathtt{T}_i^2$, and its right–invariance under $\mathcal{G}l(k)$, are readily apparent. Observe that developments including the definition of $\mathfrak{M}$ are devoid of $\boldsymbol{\delta}_i$, *i.e.* $\boldsymbol{\delta}_i = \boldsymbol{0}$, so that Lemma 2 establishes invariance of the *null* distribution $\mathcal{L}(\mathtt{T}_i^2|H_0)$. Conclusion (ii) follows on recognizing that moments are not required in the developments of earlier sections. Specifically, $\boldsymbol{\Lambda} = \boldsymbol{1}_n\cdot\boldsymbol{\mu}' + [\boldsymbol{O}',\ \boldsymbol{\delta}_i]'$ of Assumptions A may be taken to be location and shift parameters without first moments, and the earlier dispersion parameters $\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}$ serve instead as scale parameters of the distributions *in lieu of* second moments. ■

*Remark* 1. It is nothing short of *remarkable* that operating characteristics of $\mathtt{T}_i^2$ should be identical under matrix Cauchy errors as under matrix Gaussian errors. The importance of heavy–tailed distributions in economics and finance is highlighted in Ibragimov *et al.* (2015) as noted.

## 4. Case Studies

### 4.1 Overview

Calibrated data often entail *calibration curves,* direct or indirect, both injecting dependencies among the calibrated measurements; see Jensen and Ramirez (2009, 2012). These apply in the analysis of univariate data. Another venue adjusts observations directly to a common standard, as in compensating for the *tare weight* of a scale, or in assessing yield increments relative to a control yield as in Jensen (2001b). Subsequent examples fall within the latter framework, which we develop next for multivariate data amenable to Hotelling's $\mathtt{T}_i^2$ diagnostics.

In short, observations $\boldsymbol{Y}_0 = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]'$ from $\{\boldsymbol{Y}_0 = \mathbf{1}_n\boldsymbol{\mu}' + \mathfrak{E}_0\}$ are sought. For Case I the user sees $\boldsymbol{Z}_0 = [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n]'$ having $\{\boldsymbol{Z}_j = \boldsymbol{Y}_j + \mathtt{W} \in \mathbb{R}^k; 1 \le j \le n\}$ with scalar shifts $\mathtt{W} = [\,\mathtt{W}_1, \ldots, \mathtt{W}_k\,]$ often themselves random. To model this we proceed as follows: (a) Append $\boldsymbol{Y}_0^{\dagger} = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n, \mathtt{W}]'$; (b) suppose that $\mathtt{V}(\boldsymbol{Y}_0^{\dagger}) = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}$ with $N = n+1$; and (c) let $\boldsymbol{A} = [\boldsymbol{I}_n, \mathbf{1}_n]$. Then $\boldsymbol{Z}_0 = \boldsymbol{A}\boldsymbol{Y}_0^{\dagger} \in \mathbb{F}_{n\times k}$. For Case II, if rows of $\boldsymbol{Y}_0$ are to be adjusted instead against $\mathtt{W}$ as standard, then $\{\boldsymbol{Z}_j = (\boldsymbol{Y}_j - \mathtt{W}); 1 \le j \le n\}$ and $\boldsymbol{A} = [\boldsymbol{I}_n, -\mathbf{1}_n]$. To continue, in both cases we have $\mathtt{V}(\boldsymbol{Z}_0) = \boldsymbol{A}\boldsymbol{I}_N\boldsymbol{A}' \otimes \boldsymbol{\Sigma}$ from Remark A.1. In what follows we parallel steps given heretofore in working from $\boldsymbol{Y}_0$ to $\mathtt{T}_i^2$.

**Lemma 3.** *Begin with* $\boldsymbol{Z}_0$ *as constructed; rearrange as* $\boldsymbol{Z}_0 = [\boldsymbol{Z}', \boldsymbol{Z}_i]'$ *with* $\boldsymbol{Z}_i$ *as the test case; and determine* $\mathtt{T}_i^2$ *from* $\boldsymbol{Z}_0$ *as before using* $\boldsymbol{Y}_0$. *Then*

*(i) The residuals* $\boldsymbol{R}_0 = \boldsymbol{B}_n\boldsymbol{Z}_0$ *have* $\mathtt{V}(\boldsymbol{R}_0) = \boldsymbol{B}_n \otimes \boldsymbol{\Sigma}$ *independently of* $\mathtt{W}$;

*(ii)* $\mathcal{L}(\mathtt{T}_i^2) = T_k^2(u; r-1, \lambda_i)$ *independently of* $\mathtt{W}$, *with* $\lambda_i = \frac{r}{n}\boldsymbol{\delta}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_i$;

*(iii)* $\mathtt{T}_i^2$ *remains exact in level and power for all mixtures in* $\mathrm{M}_1$ *of (5);*

*(iv) These* $\mathtt{T}^2$ *distributions are identical to those initially derived from the unadjusted* $\mathcal{L}(\boldsymbol{Y}_0) = \mathrm{N}_{n\times k}(\boldsymbol{\Lambda}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$.

**Proof.** Taking $\mathtt{V}(\boldsymbol{Z}_0) = \boldsymbol{A}\boldsymbol{A}' \otimes \boldsymbol{\Sigma}$ from $\boldsymbol{A} = [\boldsymbol{I}_n, \pm\mathbf{1}_n]$, it follows that $\boldsymbol{B}_n\boldsymbol{A}\boldsymbol{A}'\boldsymbol{B}_n \otimes \boldsymbol{\Sigma} = \boldsymbol{B}_n \otimes \boldsymbol{\Sigma}$ since $\boldsymbol{B}_n$ is idempotent, to give conclusion (i). Conclusion (ii) follows directly, setting the stage so that conclusions (iii)–(iv) now follow from Theorem 3, to complete our proof. ■

*Remark* 2. Observe that a fractional adjustment $\{\boldsymbol{Z}_j = (\boldsymbol{Y}_j \pm \kappa\mathtt{W}); 1 \le j \le n\}$ can be achieved on taking $\boldsymbol{A} = [\boldsymbol{I}_n, \pm\kappa\,\mathbf{1}_n]$. The stated conclusions follow directly if so modified.

### 4.2 Simulation Studies

As developments heretofore are tedious, convoluted, and unconventional, it is instructive to demonstrate Theorem 2 and then Theorem 3. Details follow.

**(i) Induced Correlations.**

Accordingly, $N = 40{,}000$ random samples $\boldsymbol{Y}_0 \in \mathbb{F}_{n\times k}$ of size $n = 10$ and $k = 2$ were generated from $\mathrm{N}_{n\times k}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$ with rows as independent bivariate Gaussian vectors having zero means and dispersion matrix $\boldsymbol{\Sigma} = \left[\begin{smallmatrix} 1.0 & 0.8 \\ 0.8 & 1.0 \end{smallmatrix}\right]$. Correlations among rows were induced through $\boldsymbol{Y}_0 \to \boldsymbol{Z}_0 = [\boldsymbol{\Omega}(\theta)]^{\frac{1}{2}}\boldsymbol{Y}_0$ using the spectral square root, so that $\mathtt{V}(\boldsymbol{Z}_0) = \boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$ using Remark A.1. In consequence, rows of $\boldsymbol{Z}_0$ are equicorrelated with $\rho = \theta/(1+\theta)$ and, to illustrate, we vary $\rho \in [\,0.0, 0.2, 0.5, 0.8]$ with corresponding $\theta \in [\,0.00, 0.25, 1.00, 4.00]$. MINITAB was used for all the simulations.

Table 1 reports the empirical critical values for $\mathtt{T}_i^2$ corresponding to tabulated critical values $c_\alpha$ from Theorem 2(vi) such that $\mathcal{L}(\mathtt{T}_i^2) = T_k^2(u; r-1, 0)$. The row being evaluated for a potential shift is set to $i = 10$. Computations yielding $\mathtt{T}_i^2$ were undertaken for each repetition, with results as summarized in Table 1.

Table 1. Tabulated and empirical critical values for $\{\mathtt{T}_i^2 \ge c_\alpha\}$, $N = 40{,}000$ runs, for correlated $\boldsymbol{Z}_0$ with varying $\rho$ such that $V(\boldsymbol{Z}_0) = \boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$ with $\theta = \rho/(1-\rho)$, $n = 10$, and $k = 2$.

| $\alpha$ | 10% | 5% | 2.5% | 1% |
|---|---|---|---|---|
| Tabulated $c_\alpha$ | 7.45 | 10.83 | 14.95 | 21.82 |
| $\rho = 0.0$ | 7.40 | 10.77 | 14.65 | 21.84 |
| $\rho = 0.2$ | 7.40 | 10.77 | 14.65 | 21.84 |
| $\rho = 0.5$ | 7.40 | 10.77 | 14.65 | 21.84 |
| $\rho = 0.8$ | 7.40 | 10.77 | 14.65 | 21.84 |

Table 2. Tabulated and empirical powers for $\{\mathtt{T}_i^2 \geq c_\alpha\}$, $N = 40,000$, with varying shifts $\boldsymbol{\delta}' = [\delta_1, \delta_2]$ for correlated $\boldsymbol{Z}_0$ such that $V(\boldsymbol{Z}_0) = \boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$ with $\theta = 1$, $n = 10$, and $\lambda_i = \frac{r}{n}\boldsymbol{\delta}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_i$.

| $\delta_1$ | $\delta_2$ | $k$ | $r-k$ | $\lambda_i$ | Tabulated Power | Empirical Power |
|---|---|---|---|---|---|---|
| 0 | 0 | 2 | 7 | 0.000 | 0.0500 | 0.0496 |
| 1 | 1 | 2 | 7 | 1.000 | 0.1020 | 0.1036 |
| 2 | 2 | 2 | 7 | 4.000 | 0.2848 | 0.2822 |
| 3 | 3 | 2 | 7 | 9.000 | 0.5670 | 0.5647 |
| 4 | 4 | 2 | 7 | 16.000 | 0.8188 | 0.8177 |
| 5 | 5 | 2 | 7 | 25.000 | 0.9504 | 0.9499 |

Table 2 reports the empirical power for $\mathtt{T}_i^2$ with noncentrality parameter $\lambda_i = \frac{r}{n}\boldsymbol{\delta}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_i$ as in Theorem 2(vi). The corresponding tabulated power as in Theorem 2(vi) is from $\mathcal{L}(\mathtt{T}_i^2) = T_k^2(u; r-1, \lambda_i)$, where the shifts were added to the residual for row $i = 10$. Table 2 demonstrates that the powers for Hotelling's $\mathtt{T}_i^2$ diagnostics under equicorrelated data are equivalent to those tabulated for independent data. Recalling that $\mathcal{L}(\frac{r-k}{k(r-1)}\mathtt{T}_i^2) = F(u; k, r-k, \lambda_i)$, the noncentral $F$ probabilities were computed using the Keisan Online Calculator provided by the Casio Computer Co., Ltd.

**(ii) Mixture Experiments.**

To demonstrate the validity of $\mathtt{T}_i^2$ in mixture distributions as in Theorem 3, $N = 40,000$ random samples $\boldsymbol{Y}_0 \in \mathbb{F}_{n \times k}$ of size $n = 10$ and $k = 2$ were generated from $\mathrm{N}_{n \times k}(\boldsymbol{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$ having zero means and dispersion matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1.0 & 0.8 \\ 0.8 & 1.0 \end{bmatrix}$. As before, correlations among the rows were induced through $\boldsymbol{Y}_0 \to \boldsymbol{Z}_1 = [\boldsymbol{\Omega}(\theta_1)]^{\frac{1}{2}}\boldsymbol{Y}_0$ so that $\mathtt{V}(\boldsymbol{Z}_1) = \boldsymbol{\Omega}(\theta_1) \otimes \boldsymbol{\Sigma}$ using Remark A.1. This was repeated with $\boldsymbol{Y}_0 \to \boldsymbol{Z}_2 = [\boldsymbol{\Omega}(\theta_2)]^{\frac{1}{2}}\boldsymbol{Y}_0$ to form another correlated data set.

For a 50%–50% mixture, $m = 5$ observations were randomly chosen from each of $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ and stacked in random order to form a data set of order $(10 \times 2)$. Table 3 reports the empirical critical values for $\mathtt{T}_i^2$ from Theorem 2(v), with the corresponding tabulated critical values $c_\alpha$ from Theorem 2(vi) where $\mathcal{L}(\mathtt{T}_i^2) = T_k^2(u; r-1, 0)$. Values $\rho \in [\,0.0, 0.2, 0.5\,]$ were used with $\theta = \rho/(1-\rho)$. The row being evaluated for a potential shift was set to be $i = 10$.

Table 3. Tabulated and empirical critical values for $\{\mathtt{T}_i^2 \geq c_\alpha\}$, $N = 40,000$, with varying $\rho$, for 50%-50% mixtures with correlated $\boldsymbol{Z}_1$ having $\mathtt{V}(\boldsymbol{Z}_1) = \boldsymbol{\Omega}(\theta_1) \otimes \boldsymbol{\Sigma}$, and with $\boldsymbol{Z}_2$ having $\mathtt{V}(\boldsymbol{Z}_2) = \boldsymbol{\Omega}(\theta_2) \otimes \boldsymbol{\Sigma}$, where $\theta = \rho/(1-\rho)$, $m = 5$, $k = 2$, and $n = 10$.

| $\boldsymbol{\Omega}(\theta_1)$ | $\boldsymbol{\Omega}(\theta_2)$ | 10% | 5% | 2.5% | 1% |
|---|---|---|---|---|---|
| $\rho_1$=0.0 | $\rho_2$=0.0 | 7.40 | 10.77 | 14.65 | 21.84 |
| $\rho_1$=0.0 | $\rho_2$=0.2 | 7.50 | 10.83 | 15.07 | 21.90 |
| $\rho_1$=0.0 | $\rho_2$=0.5 | 7.31 | 10.51 | 14.55 | 21.41 |
| $\rho_1$=0.2 | $\rho_2$=0.5 | 7.21 | 10.29 | 14.27 | 20.90 |
| Tabulated $c_\alpha$ | | 7.45 | 10.83 | 14.95 | 21.82 |

Table 3 demonstrates empirically the invariance of $\mathtt{T}_i^2$ in mixture experiments. Observe that the second and third cases comprise contamination of $\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}$, the classical model, with 50% contamination using $\boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$.

*4.3 Running Times Example*

Woodward (1970) studied the running times for $n = 22$ baseball players who ran three different paths rounding first base. These data, as used by Morrison (2005) to test for outliers, are reported in Appendix Table 7. The times that appear to be abnormal are those for Player 14 and Player 22. Using $\mathtt{T}_i^2$ we see in Table 4 that the running times for Player 22 are indeed outlying with $p$–value 0.0138. Times for Player 14 are not flagged as outliers. However, the last four rows of Table 4 give $\mathtt{T}_i^2$ for Player 14 assuming improvements in his running times in units of $\delta \in [\,0.1, 0.2, 0.3, 0.4\,]$.

Beckett (1977) has identified that the $n = 22$ data points consist of two clusters, namely [2, 4, 5, 7–15, 17, 19–22] and [1, 3, 6, 16, 18], where each cluster consists of correlated data. Observe that the rank order of the times $[Y_1, Y_2, Y_3]$ for Players 1, 6, and 18 differ from the rank order for other players. In consequence, Morrison's (2005) search for outliers using $\mathtt{T}_i^2$ is in dispute when based on the validating model of the day, namely $\mathtt{V}(\boldsymbol{Y}) = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}$, instead of the apparent mixing over clusters.

A fortunate conclusion from Theorems 2 and 3 is that searches for outliers using $\mathtt{T}_i^2$ now has been validated for data sets from cohorts which are equicorrelated. And, in addition, even for data sets arising as mixture distributions, in a belated

Table 4. Values of $\mathtt{T}_i^2$ for Player 22 and Player 14 with varying improvements $\boldsymbol{\delta}' = [\delta_1, \delta_2, \delta_3]$ for Player 14.

| Player | $\boldsymbol{\delta} = [\delta_1, \delta_2, \delta_3]'$ | $\mathtt{T}_i^2$ | $n$ | $k$ | $r-1$ | $p$-value |
|---|---|---|---|---|---|---|
| 22 | | 15.589 | 22 | 3 | 20 | 0.0138 |
| 14 | | 4.886 | 22 | 3 | 20 | 0.2572 |
| 14 | $(-0.1, -0.1, -0.1)$ | 6.835 | 22 | 3 | 20 | 0.1432 |
| 14 | $(-0.2, -0.2, -0.2)$ | 9.094 | 22 | 3 | 20 | 0.0744 |
| 14 | $(-0.3, -0.3, -0.3)$ | 11.694 | 22 | 3 | 20 | 0.0367 |
| 14 | $(-0.4, -0.4, -0.4)$ | 14.624 | 22 | 3 | 20 | 0.0175 |

verification of Morrison's (2005) analysis.

### *4.4 Adjusting to a Common Standard*

Following the notation of Section 4.1, we consider a Case II example in which the data $\boldsymbol{Y}_0 = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n]'$ are adjusted against a standard $\mathtt{W} = [\mathtt{W}_1, \ldots, \mathtt{W}_k]$. Here $N = 40,000$ random samples $\boldsymbol{Y}_0 \in \mathbb{F}_{n\times k}$ were generated from $N_{n\times k}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}(\rho))$ of size $n = 20$ and $k = 4$ with rows as independent Gaussian vectors having zero means and dispersion matrix $\boldsymbol{\Sigma}(\rho) = (1-\rho)\boldsymbol{I}_4 + \rho\mathbf{1}_4\mathbf{1}_4'$ with $\rho = 0.8$. The standardizing vector $\mathtt{W}$ was random from $N_k(\mathbf{0}, \boldsymbol{\Sigma})$, with $\{\boldsymbol{Z}_j = \boldsymbol{Y}_j - \mathtt{W}; 1 \le j \le n\}$ and with $\boldsymbol{Z}_0 = \boldsymbol{A}\boldsymbol{Y}_0^{\dagger}$ as equicorrelated data with $\mathtt{V}(\boldsymbol{Z}_0) = \boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$, $\theta = 1.0$ and corresponding $\rho = \theta/(1 + \theta) = 0.5$. Lemma 3(iii) notes that $\mathtt{T}_i^2 = (r-1)\mathrm{tr}(\boldsymbol{Q}_1\boldsymbol{Q}_2^{-1})$ remains exact in level and power for the equicorrelated data. Table 5 reports the tabulated and empirical critical values for this study, affirming that the critical values remain the same for data equicorrelated by adjustment to a common standard.

Table 5. Tabulated and empirical critical values for $\{\mathtt{T}_i^2 \ge c_\alpha\}$, $N = 40,000$ runs, for correlated $\{\boldsymbol{Z}_j = \boldsymbol{Y}_j - \mathtt{W}\}$ as $\boldsymbol{Y}_j$ adjusted to the standard $\mathtt{W}$, with $\mathtt{V}(\boldsymbol{Z}_0) = \boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$ and $\theta = 1.0$, $\rho = 0.5$, $n = 20$ and $k = 4$.

| $\alpha$ | 10% | 5% | 2.5% | 1% |
|---|---|---|---|---|
| Tabulated $c_\alpha$ | 11.33 | 14.67 | 18.26 | 23.49 |
| Empirical $c_\alpha, \rho = 0.5$ | 11.32 | 14.70 | 18.24 | 23.38 |

Table 6. Tabulated and empirical powers for $\{\mathtt{T}_i^2 \ge c_{0.05}\}$, $N = 40,000$ runs, with varying shifts $\boldsymbol{\delta}' = [\delta_i, \delta_i, \delta_i, \delta_i]$ for correlated $\{\boldsymbol{Z}_j = \boldsymbol{Y}_j - \mathtt{W}\}$ as $\boldsymbol{Y}_j$ adjusted to the standard $\mathtt{W}$, with $\mathtt{V}(\boldsymbol{Z}_0) = \boldsymbol{\Omega}(\theta) \otimes \boldsymbol{\Sigma}$ and $\theta = 1.0$, $\rho = 0.5$, $n = 20$ and $k = 4$, and $\lambda_i = \frac{r}{n}\boldsymbol{\delta}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_i$.

| $\delta_i$ | $k$ | $r-k$ | $\lambda_i$ | Tabulated Power | Empirical Power |
|---|---|---|---|---|---|
| 0 | 4 | 15 | 0.000 | 0.0500 | 0.0503 |
| 1 | 4 | 15 | 1.118 | 0.0956 | 0.0933 |
| 2 | 4 | 15 | 4.471 | 0.2683 | 0.2681 |
| 3 | 4 | 15 | 10.059 | 0.5673 | 0.5694 |
| 4 | 4 | 15 | 17.882 | 0.8391 | 0.8388 |
| 5 | 4 | 15 | 27.941 | 0.9661 | 0.9661 |

Table 6 reports the tabulated and empirical powers for $\{\mathtt{T}_i^2 \ge c_\alpha\}$ with $c_{0.05} = 14.67$ and with non-centrality parameter $\lambda_i = \frac{r}{n}\boldsymbol{\delta}_i'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_i$. The shifts $\boldsymbol{\delta}_i' = [\delta_1, \ldots, \delta_4]$ used a common value $\delta_i$ and were added to the residuals for row $i = 20$.

## 5. Conclusions

The *objectives* set forth in the Introduction now have been met, namely, the construction of dispersion mixtures of matrix distributions under direct products of the type $\mathtt{V}(\boldsymbol{Y}_0) = \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}$, together with the invariance of null and nonnull distributions of $\mathtt{T}_i^2$ under such mixtures. Additional findings establish the null distribution of $\mathtt{T}_i^2$ to be invariant under left–spherical symmetry, encompassing matrix stable and Cauchy distributions as special cases. These findings do expand considerably the theoretical bases validating the use of $\mathtt{T}_i^2$ in practice.

To place this study in perspective, users long have recognized that classical Gaussian models often are inadequate under the exigencies of contemporary research. Early remedial efforts focused on spherical and elliptical symmetries in $\mathbb{R}^n$, generating many well known research papers and monographs. More recent studies examine additional structural properties in $\mathbb{R}^n$; examples include Arnold *et al.* (2008), Kamiya *et al.* (2008), Sarabia and Gómez-Déniz (2008), Richter (2013), and Richter (2014). These include mixtures, asymmetries, and various star–contoured densities in $\mathbb{R}^n$. Among the latter are the dispersion mixtures of Jensen and Ramirez (2015), having the remarkable feature that the distributions of $t_i^2$ and those of Dixon (1950), Grubbs (1950), and Ferguson (1961) are all invariant and thus identical to their normal theory forms.

The present study breaks new ground in extending the *n*–dimensional mixture distributions of Jensen and Ramirez (2015) to include star–contoured matrix distributions in $\mathbb{F}_{n\times k}$, together with invariance of the distribution of $\mathtt{T}_i^2$. The case studies offer further insight regarding the extended uses of $\mathtt{T}_i^2$ in practice.

## A Appendix

### *A.1 Matrix Distributions*

We collect basics for matrix distributions essential to the present study. First partition $\boldsymbol{Y}_0$ by columns as $\boldsymbol{Y}_0=[\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_k]\in\mathbb{F}_{n\times k}$. Alternatively, if instead we row–partition as $\boldsymbol{Y}_0'=[\boldsymbol{Z}_1,\ldots,\boldsymbol{Z}_n]$, then we adopt the convention that $\mathtt{V}(\boldsymbol{Y}_0)=\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}$, so that $\{\mathtt{V}(\boldsymbol{Z}_i)=\omega_{ii}\boldsymbol{\Sigma},\ 1\le i\le n\}$. Moreover, if $\mathtt{V}(\boldsymbol{Y})=\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}$ of order $(nk\times nk)$, then for fixed $(\boldsymbol{A},\boldsymbol{B})$ and for $\boldsymbol{U}=\boldsymbol{AYB}'$, the corresponding moment arrays are as follow.

*Remark* 3. $\mathtt{E}(\boldsymbol{Z})=\boldsymbol{AMB}'$ and $\mathtt{V}(\boldsymbol{Z})=\boldsymbol{A\Omega A}'\otimes\boldsymbol{B\Sigma B}'$.

(i) Gaussian Distributions

Designate the distribution of $\boldsymbol{Y}\in\mathbb{F}_{n\times k}$ as $\mathcal{L}(\boldsymbol{Y})=\mathrm{N}_{n\times k}(\boldsymbol{M},\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$, the *matrix Gaussian* distribution in $\mathbb{F}_{n\times k}$ having $\mathtt{E}(\boldsymbol{Y})=\boldsymbol{M}$ and $\mathtt{V}(\boldsymbol{Y})=\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}$. Its *pdf* is

$$f_{nk}(\boldsymbol{Y})=(2\pi)^{-\frac{nk}{2}}|\boldsymbol{\Omega}|^{-\frac{k}{2}}|\boldsymbol{\Sigma}|^{-\frac{n}{2}}\exp\left[-\frac{1}{2}\,\mathrm{tr}\,\boldsymbol{\Omega}^{-1}(\boldsymbol{Y}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{M})'\right]. \tag{13}$$

Next partition $\boldsymbol{Y}=[\boldsymbol{Y}_1,\boldsymbol{Y}_2]$, $\boldsymbol{M}=[\boldsymbol{M}_1,\boldsymbol{M}_2]$, and $\boldsymbol{\Sigma}=[\boldsymbol{\Sigma}_{ij};i,j=1,2]$. Here $(\boldsymbol{Y}_1,\boldsymbol{M}_1)$ are of order $(n\times r)$; $(\boldsymbol{Y}_2,\boldsymbol{M}_2)$ are of order $(n\times s)$; and elements of $\boldsymbol{\Sigma}$ are $\{\boldsymbol{\Sigma}_{11}(r\times r),\ \boldsymbol{\sigma}_{12}(r\times s),\ \boldsymbol{\sigma}_{21}(s\times r)\},\ \boldsymbol{\sigma}_{22}(s\times s)\}$. Then

- Marginals: $\mathcal{L}(\boldsymbol{Y}_1)=\mathrm{N}_{n\times r}(\boldsymbol{M}_1,\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}_{11})$; $\mathcal{L}(\boldsymbol{Y}_2)=\mathrm{N}_{n\times s}(\boldsymbol{M}_2,\boldsymbol{\Omega}\otimes\boldsymbol{\sigma}_{22})$.
- Conditional: $\mathcal{L}(\boldsymbol{Y}_1|\boldsymbol{Y}_2=\boldsymbol{y}_2)=\mathrm{N}_{n\times r}(\boldsymbol{M}_{1\cdot2},\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}_{11\cdot2})$ with $\boldsymbol{M}_{1\cdot2}=\boldsymbol{M}_1+(\boldsymbol{y}_2-\boldsymbol{M}_2)\boldsymbol{R}'$ and $\boldsymbol{R}=\boldsymbol{\sigma}_{12}\boldsymbol{\sigma}_{22}^{-1}$.

(ii) Wishart Distributions

Take $\mathcal{L}(\boldsymbol{Y})=\mathrm{N}_{n\times k}(\boldsymbol{M},\boldsymbol{I}_n\otimes\boldsymbol{\Sigma})$; let $\boldsymbol{W}=\boldsymbol{Y}'\boldsymbol{Y}$; and partition $\boldsymbol{W}=[\boldsymbol{W}_{ij};i,j=1,2]$ and $\boldsymbol{\Sigma}=[\boldsymbol{\Sigma}_{ij};i,j=1,2]$ conformably. Here elements of $\boldsymbol{W}$ are $\{\boldsymbol{W}_{11}(r\times r),\ \boldsymbol{W}_{12}(r\times s),\ \boldsymbol{W}_{21}(s\times r),\ \boldsymbol{W}_{22}(s\times s)\}$. Define $\boldsymbol{W}_{11\cdot2}=\boldsymbol{W}_{11}-\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{21}$, and $\boldsymbol{\Sigma}_{11\cdot2}$ as before.

- $\boldsymbol{W}$ is said to have the Wishart distribution $\mathcal{L}(\boldsymbol{W})=\mathbb{W}_k(n;\boldsymbol{\Sigma},\boldsymbol{\Theta})$ with $n$ degrees of freedom, the scale parameters $\boldsymbol{\Sigma}$, and the noncentrality matrix $\boldsymbol{\Theta}=\boldsymbol{M}'\boldsymbol{M}$.
- Marginals: $\mathcal{L}(\boldsymbol{W}_{11})=\wr(n;\boldsymbol{\Sigma}_{11},\boldsymbol{M}_1'\boldsymbol{M}_1)$ and $\mathcal{L}(\boldsymbol{W}_{22})=\mathrm{W}_s(n,\boldsymbol{\sigma}_{22},\boldsymbol{M}_2'\boldsymbol{M}_2)$.
- Conditional: $\mathcal{L}(\boldsymbol{W}_{11\cdot2}|\boldsymbol{Y}_2=\boldsymbol{y}_2)=\wr(n-s;\boldsymbol{\Sigma}_{11\cdot2},\boldsymbol{\Theta}(\boldsymbol{y}_2))$ with $\boldsymbol{\Theta}(\boldsymbol{y}_2)=\boldsymbol{M}_{1\cdot2}'\boldsymbol{M}_{1\cdot2}$ and $\boldsymbol{M}_{1\cdot2}=\boldsymbol{M}_1+\boldsymbol{R}(\boldsymbol{y}_2-\boldsymbol{M}_2)$ with$\boldsymbol{R}=\boldsymbol{\sigma}_{12}\boldsymbol{\sigma}_{22}^{-1}$.
- A standard result is that if $\boldsymbol{A}$ is idempotent of rank $\nu$, and if $\boldsymbol{W}=\boldsymbol{Y}'\boldsymbol{AY}$, then $\mathcal{L}(\boldsymbol{W})=\mathbb{W}_k(\nu,\boldsymbol{\Sigma},\boldsymbol{\Theta})$ with $\boldsymbol{\Theta}=\boldsymbol{M}'\boldsymbol{AM}$.

(iii) A characterization

Fundamental connections link noncentral chi–squared and noncentral Wishart distributions. To wit: The noncentral Chi–squared and Wishart *chf*'s are

$$\phi_Z(t)=(1-2it\sigma^2)^{-\frac{n}{2}}\exp[\,it\,\theta^2/(1-2it\sigma^2)]$$
$$\phi_W(\boldsymbol{T})=|\boldsymbol{I}_k-2i\,\boldsymbol{T\Sigma}|^{-\frac{n}{2}}\exp[\,i\,\mathrm{tr}\,\boldsymbol{T\Theta}(\boldsymbol{I}_k-2i\,\boldsymbol{T\Sigma})^{-1}],$$

respectively. The following is germane.

**Theorem 5.** *Let $\mathcal{L}(\boldsymbol{Y})=\mathrm{N}_{n\times k}(\boldsymbol{M},\boldsymbol{I}_n\otimes\boldsymbol{\Sigma})$, and take $\boldsymbol{W}=\boldsymbol{Y}'\boldsymbol{AY}$ such that $\boldsymbol{A}$ is idempotent of rank $\nu$.*

*(i) If $\mathcal{L}(\boldsymbol{W})=\mathbb{W}_k(\nu,\boldsymbol{\Sigma},\boldsymbol{\Theta})$, then $\mathcal{L}(\boldsymbol{u}'\boldsymbol{Wu})=\chi^2(\nu,\sigma_{\boldsymbol{u}}^2,\lambda(\boldsymbol{u}))$ for every $\boldsymbol{u}\in\mathbb{R}^k$, where $\sigma_{\boldsymbol{u}}^2=\boldsymbol{u}'\boldsymbol{\Sigma u}$ and $\lambda(\boldsymbol{u})=\boldsymbol{u}'\boldsymbol{\Theta u}$.*

*(ii) Conversely, if* $\mathcal{L}(\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}) = \chi^2(\nu, \sigma^2_{\boldsymbol{u}}, \lambda(\boldsymbol{u}))$ *for every* $\boldsymbol{u} \in \mathbb{R}^k$, *then* $\mathcal{L}(\boldsymbol{W}) = \mathbb{W}_k(\nu, \boldsymbol{\Sigma}, \boldsymbol{\Theta})$.

*(iii) Define* $\boldsymbol{W}_1 = \boldsymbol{Y}'\boldsymbol{A}_1\boldsymbol{Y}$ *and* $\boldsymbol{W}_2 = \boldsymbol{Y}'\boldsymbol{A}_2\boldsymbol{Y}$ *with* $\boldsymbol{A}_1 \neq \boldsymbol{A}_2$. *Then* $(\boldsymbol{W}_1, \boldsymbol{W}_2)$ *are mutually independent Wishart matrices if and only* $(\boldsymbol{u}'\boldsymbol{W}_1\boldsymbol{u},\ \boldsymbol{u}'\boldsymbol{W}_2\boldsymbol{u})$ *are mutually independent* $\chi^2$ *variates for every* $\boldsymbol{u} \in \mathbb{R}^k$.

**Proof.** Conclusion (i) follows on substituting $t\boldsymbol{u}\boldsymbol{u}'$ for $\boldsymbol{T}$ in the *chf* for $\boldsymbol{W}$, together with the fact that the nonvanishing eigenvalue of $\boldsymbol{u}\boldsymbol{u}'$ is $\boldsymbol{u}'\boldsymbol{u}$. Conclusion (ii) follows on lifting from $\mathbb{R}^1_+$ to $\mathbb{S}^+_k$; this may be done using the characterization of Cramér and Wold (1936), as carried out in Jensen (1982). Conclusion (iii) follows from (i) and (ii) on verifying that the joint *chf*'s of $(\boldsymbol{W}_1, \boldsymbol{W}_2)$ and of $(\boldsymbol{u}'\boldsymbol{W}_1\boldsymbol{u},\ \boldsymbol{u}'\boldsymbol{W}_2\boldsymbol{u})$ factor into the product of their marginal *chf*'s. ■

*Remark* 4. The central version of conclusions (i) and (ii) was given in Result (ii) of Rao (1973, page 535).

*A.2 Running Times Data*

The data employed in Section 4.3 are listed here as reported in Morrison (2005, page 102).

Table 7. Running times around first base for $k = 3$ paths and $n = 22$ players.

| Player | $Y_1$ | $Y_2$ | $Y_3$ | Player | $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|---|---|---|---|---|
| 1 | 5.40 | 5.50 | 5.55 | 12 | 5.65 | 5.55 | 5.45 |
| 2 | 5.85 | 5.70 | 5.75 | 13 | 5.60 | 5.35 | 5.45 |
| 3 | 5.20 | 5.60 | 5.50 | 14 | 5.05 | 5.00 | 4.95 |
| 4 | 5.55 | 5.50 | 5.40 | 15 | 5.50 | 5.50 | 5.40 |
| 5 | 5.90 | 5.85 | 5.70 | 16 | 5.45 | 5.55 | 5.50 |
| 6 | 5.45 | 5.55 | 5.60 | 17 | 5.55 | 5.55 | 5.35 |
| 7 | 5.40 | 5.40 | 5.35 | 18 | 5.45 | 5.50 | 5.55 |
| 8 | 5.45 | 5.50 | 5.35 | 19 | 5.50 | 5.45 | 5.25 |
| 9 | 5.25 | 5.05 | 5.00 | 20 | 5.65 | 5.60 | 5.40 |
| 10 | 5.85 | 5.80 | 5.70 | 21 | 5.70 | 5.65 | 5.55 |
| 11 | 5.25 | 5.20 | 5.10 | 22 | 6.30 | 6.30 | 6.25 |

## References

Aitkin, M., & Wilson, G. (1980). Mixture models, outliers, and the EM algorithm. *Technometrics, 22*, 325–331. https://doi.org/10.1080/00401706.1980.10486163

Arnold, B. C., Castillo, E., & Sarabia, J. M. (2008). Multivariate distributions defined in terms of contours. *Journal of Statistical Planning and Inference, 138*, 4158–4171. https://doi.org/10.1016/j.jspi.2008.03.033

Barrett, B. E. (2003). Understanding influence in multivariate regression. *Commununications in Statistics: Theory and Methods, 32*, 667–680. https://doi.org/10.1081/STA-120018557

Barrett, B. E., & Ling, R. F. (1992). General classes of influence measures for multivariate regression. *Journal of the American Statistical Association, 87*, 184–191. https://doi.org/ 10.1080/01621459.1992.10475191

Beckett, J. (1977). Clustering rank preference data, in ASA Proceedings of the Social Statistics Section, 983–986.

Beckman. R. J., & Trussell H. J. (1974). The distribution of an arbitrary Studentized residual and the effects of updating in multiple regression. *Journal of the American Statistical Association, 69*, 199–201. https://doi.org/10.1080/01621459.1974.10480152

Box, G. E. P., & Tiao, G. C. (1968). A Bayesian approach to some outlier problems. *Biometrika, 55*, 119–129. https://doi.org/10.1093/biomet/55.1.119

Caroni, C. (1987). Residuals and influence in the multivariate linear model. *Journal of the Royal Statistical Society Series D, 36*, 365–370. https://doi.org/10.2307/2348833

Cramér, H., & Wold, H. (1936). Some theorems on distribution functions. *Journal of the London Mathematical Society, 11*, 290–294. https://doi.org/10.1112/jlms/s1-11.4.290

Dixon, W. J. (1950). Analysis of extreme values. *Annals of Mathematical Statistics, 21*, 488–506. https://doi.org/10.1214/aoms/1177729747

Ferguson, T. S. (1961). On the rejection of outliers. *In Proceedings of the 4th Berkeley Symposium, 1*, 253–287. Statistics, vol 214. Berlin: Springer–Verlag.

Grubbs, F. E. (1950). Criteria for testing outlying observations. *Annals of Mathematical Statistics, 21*, 27–58.

Hotelling, H. (1931). The generalization of Student's ratio. *Annals of Mathematical Statistics, 2*, 360–378. https://doi.org/10.1214/aoms/1177732979

Ibragimov, M., Ibragimov, R., & Walden, J. (2015). Implications of heavy-tailedness. In: Heavy-Tailed Distributions and Robustness in Economics and Finance. *Lecture Notes in Statistics, 214*. Berlin: Springer–Verlag.

Jensen, D. R. (1982). Recovering the Wishart distribution from its sections. *Journal of Statistical Computation and Simulation, 15*, 69–72. https://doi.org/10.1080/00949658208810564

Jensen, D. R. (1996). Structured dispersion and validity in linear inference. *Linear Algebra and Its Applications, 249*, 189–196. https://doi.org/10.1016/0024-3795(95)00354-1

Jensen, D. R. (2001a). Properties of selected subset diagnostics in regression. *Statistics and Probability Letters, 51*, 377–388. https://doi.org/10.1016/S0167-7152(00)00178-4

Jensen, D. R. (2001b). Recovering information from incremental responses. *Australian New Zealand Journal of Statistics, 43*, 455–459. https://doi.org/10.1111/1467-842X.00193

Jensen, D. R., & Good, I. J. (1981). Invariant distributions associated with matrix laws under structural symmetry. *Journal of the Royal Statistical Society Series B, 43*, 327–332. http://www.jstor.org/stable/2984944

Jensen, D. R., & Ramirez, D. E. (2009). Anomalies in the analysis of calibrated data. *Journal of Statistical Computation and Simulation, 79*, 299–314. https://doi.org/10.1080/00949650701756864

Jensen, D. R., & Ramirez, D. E. (2012). Irregularities in X(Y) from Y(X) in linear regression. *Journal of Statistical Computation and Simulation, 83*, 1807–1828. https://doi.org/10.1080/00949655.2012.671324

Jensen, D. R., & Ramirez, D. E. (2014). Noncentralities induced in regression diagnostics. *Journal of Statistical Theory and Practice, 8*, 1–25. https://doi.org/10.1080/15598608.2014.847758

Jensen, D. R., & Ramirez, D. E. (2015). Outlier detection under star–contoured errors. *Journal of Statistical Theory and Practice, 9*, 850-867. https://doi.org/10.1080/15598608.2015.1034386

Kamiya, H, Takemura, A, & Kuriki, S. (2008). Star-shaped distributions and their generalizations. *Journal of Statistical Planning and Inference, 138*, 3429-3447. https://doi.org/10.1016/j.jspi.2006.03.016

Mathai, A. M., & Provost, S. B. (1992). Quadratic Forms in Random Variables: Theory and Applications. New York: Marcel Dekker.

Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Science of India, 12*, 49–55.

Morrison, D. F. (2005). Multivariate Statistical Methods, Fourth ed. Belmont, CA: Brooks/Cole Thomson Learning.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications, Second ed. New York: John Wiley and Sons. https://doi.org/10.1002/9780470316436

Richter, W. D. (2013): Geometric and stochastic representations for elliptically contoured distributions. *Communications in Statistics: Theory & Methods. 42*, 579-602. https://doi.org/10.1080/03610926.2011.611320

Richter, W. D. (2014). Geometric disintegration and star-shaped distributions. *Journal of Statistical Distributions and Applications, 1*(20), 1–24. https://doi.org/10.1186/s40488-014-0020-6

Rodrigues, I. M., & Boente, G. (2011). Multivariate Outliers. *In International Encyclopedia of Statistical Science, M. Lovric, ed. Part 13*, 910–912. Berlin: Springer–Verlag. https://doi.org/10.1007/978-3-642-04898-2_396

Sarabia, J. M., & Gómez-Déniz, E. (2008). Construction of multivariate distributions: a review of some recent results. *SORT, 32*, 3–36. https://www.idescat.cat/sort/sort321/32.1.1.sarabia-gomez.pdf

Snedecor, G. W., & Cochran, W. G. (1968). Statistical Methods, Sixth ed. Ames, IA: Iowa State University Press.

Woodward, W. F. (1970). A Comparison of Base Running Methods in Baseball, M.S. Thesis, Florida State University.

**Copyrights**