Research Article
Open Access

Location and scale under exchangeable errors

Abstract

Classical multivariate analysis rests on observations $Y(n\times k)$ having $n>k$ mutually independent rows, with dispersion matrix as a direct product $V(Y)=I_n\otimes\Sigma$, supported in turn by a rich literature. That independence may fail is modeled here on taking the rows of $\mathbf{Y}$ to be exchangeably dependent such that $V(Y)=\Omega\otimes\Sigma$ where exchangeability rests on the choice for $\Omega(n\times n)$. Three choices are considered; each interjects additional parameters into the model; and it remains to ask which, if any, of findings widely known under independence, might apply also under exchangeable dependence. Conventional inferences for the location and scale parameters (μ,Σ) are reconsidered. Excluding μ these are found to carry over in large part to include the exchangeable errors of this study.

AMS subject classification: 62E15, 62H15, 62J20

Keywords: exchangeable matrix errors, recovered properties, location and scale parameters

Volume 12 Issue 3 - 2023

DR Jensen
Department of Statistics, Virginia Tech, USA

Correspondence: DR Jensen, Department of Statistics, Virginia Tech, USA, Email djense@vt.edu

Received: May 23, 2023 | **Published:** June 29, 2023

Introduction

The model $\mathcal{M}\equiv\{Y=1_n\mu+\mathcal{E}\}$ asserts the n rows of $Y(n\times k)$ to be k–dimensional responses, having location parameters $\mu=[\mu_1,\mu_2,...\mu_k]$ where $1_n=[1,1,....1]'$ and with $\varepsilon(n\times k)$ as an array of random errors. Conventional analysts take the n rows of $\mathbf{Y}$ to be mutually independent and Gaussian, so that $V(Y)=I_n\otimes\Sigma$. To the contrary, independence often fails; venues include multiple time series, econometrics, and empirical adjustments that induce dependencies among the adjusted responses, as in references[1,2] for calibrated data. Accordingly, it is instructive to replace independence among rows by exchangeable dependence, on letting $V(Y)=\Omega\otimes\Sigma$ where exchangeably rests on the choice for $\Omega(n\times n)$.

In short, the basic foundations remain to be reworked, as in this study with regard to independence. Specifically, with $\mathbb{R}^n$ as Euclidean n –space and $\mathbb{F}_{n\times k}$ the real matrices of order $(n\times k)$ then the distribution $\mathcal{L}(\mathbf{Y})$ for $\mathbf{Y}\in\mathbb{R}^n$ is said to be exchangeable provided that $\mathcal{L}(P_n\boldsymbol{Y})=\mathcal{L}(\boldsymbol{Y})$ for every $P_n\in\mathbb{P}_n$ the $(n\times n)$ permutation group, a concept due to Johnson.[3] In this study exchangeable errors on $\mathbb{F}_{n\times k}$ are identified; their use is seen to offer a rich class of alternatives to independence. A brief survey follows.

Selected classes of exchangeable errors on $\mathbb{F}_{n\times k}$ are studied, as are moments for the model M. The focus here centers on (μ,Σ) as the conventional location/scale parameters. But since additional parameters are injected into the model on requiring that it should be exchangeable, it is essential to identify those properties, if any, which do carry over to include exchangeable errors.

Preliminaries

Notation

Identify $\mathbb{R}^n$ and $\mathbb{F}_{n\times k}$ as stated, with $\mathbb{S}_n^+$ as the symmetric, positive definite matrices of order $(n\times n)$. Vectors and matrices are in bold type, with $\{\boldsymbol{A}',\boldsymbol{A}^{-1},\mathrm{tr}(\boldsymbol{A}), and\,|\boldsymbol{A}|\}$ as the transpose, inverse, trace, and determinant of $\boldsymbol{A}$. The unit vector in $\mathbb{R}^n$ is $1_n=[1,1,....1]'$; $\boldsymbol{I}_n$ is the $(n\times n)$ identity; $\boldsymbol{J}_n=1_n1_n'$ and $\mathrm{Diag}(\boldsymbol{A}_1...,\boldsymbol{A}_k)$ is block–diagonal. Take $\mathrm{Ch}(\boldsymbol{A})=[\alpha_1\geq...\geq\alpha_n>0]$ to be the eigenvalues of $\boldsymbol{A}\in\mathbb{S}_n^+$. The condition number of $\boldsymbol{A}\in\mathbb{S}_n^+$ is $\mathrm{Cnd}(\boldsymbol{A})=\alpha_1/\alpha_n$. For $\boldsymbol{A}(n\times n)$ and $\boldsymbol{B}(k\times k)$, their direct product is $\boldsymbol{A}\otimes\boldsymbol{B}=[a_{ij}\boldsymbol{B}]$ of order $(nk\times nk)$, and a g –inverse of $\boldsymbol{A}\in\mathbb{F}_{n\times k}$ is $\boldsymbol{A}^-\in\mathbb{F}_{k\times n}$ such that $\mathbf{AA^-A=A}$.

Random arrays

Consider $\mathbf{Y}\in\mathbb{F}_{n\times k}$ to be random, with $\{\mathcal{L}(\mathbf{Y}),E(\mathbf{Y}),V(\mathbf{Y})\}$ as its law of distribution, its expected values in $\mathbb{F}_{n\times k}$, and its dispersion matrix in $\mathbb{S}_{nk}^+$ under moments of first and second orders. Moreover, for displaying the elements of $V(\mathbf{Y})=X_i(nk\times nk)$, the matrix $\mathbf{Y}=[\mathbf{Y}_1,\mathbf{Y}_2,...,\mathbf{Y}_n]'$ of order $(n\times k)$ is taken row–wise through the mapping $J(\mathbf{Y})=[\mathbf{Y}_1',\mathbf{Y}_2'...,\mathbf{Y}_n']$ of order $(nk\times 1)$ as in the following from Jensen DR, et al.[4]

Proposition 1:

i. For $\boldsymbol{Y}\in\mathbb{F}_{n\times k}$, then $V(\boldsymbol{Y})$ is arrayed as $V(\mathbf{Y})=\boldsymbol{X}_i$, often of the form $\boldsymbol{X}_i=\Omega\otimes\Sigma$ with elements $\{Cov(Y_i,Y_j)=\omega_{ij}\Sigma\}$;

ii. Then for row $\mathbf{Y}_i$ of $\mathbf{Y}$ the element $\{V(Y_i)=\omega_{ii}\boldsymbol{\Sigma}\}$ is on the diagonal of $\boldsymbol{X}_i$, and $\{Cov(Y_i,Y_j)=\omega_{ij}\boldsymbol{\Sigma}\}$ is off the diagonal;

iii. For $V(Z)=\Omega\otimes\Sigma$ and fixed $(\boldsymbol{A},\boldsymbol{B})$, then $V(A'ZB)=A'\Omega A\otimes B'\Sigma B$.

Exchangeable arrays trace to Johnson[3] as noted, and since have a rich history. Any mixture of independent, identically distributed (iid) variables in $\mathbb{R}^n$ is exchangeable; a converse of Finetti B[5] is that elements of $\{Y_1,Y_2,Y_3,...\}$ if exchangeable, are conditionally iid given some $Z\in\mathbb{R}^1$. Matrix arrays are considered next; refer also to Aldous.[6]

Definition 1: The distribution of $\boldsymbol{Y}\in\mathbb{F}_{n\times k}$ is said to be left–exchangeable provided that $\mathcal{L}(\boldsymbol{Y})=\mathcal{L}(P_n'\boldsymbol{Y})$ for every $\mathbf{P}_n\in\mathbb{P}_n$;

$\mathcal{L}(\boldsymbol{Y})$ is right–exchangeable provided that $\mathcal{L}(\mathbf{Y})=\mathcal{L}(\mathbf{Y}\boldsymbol{Q}_k)$ for every $\boldsymbol{Q}_k\in\mathbb{P}_k$.

Essential properties may be listed as follow.

Lemma 1: Take $\boldsymbol{y}\in\mathbb{R}^n$ with $V(y)=\Omega$, and $\boldsymbol{Y}\in\mathbb{F}_{n\times k}$ with $V(\mathrm{Y})=\Omega\otimes\Sigma$.

i. Let $\mathcal{L}(\boldsymbol{y})$ be exchangeable on $\mathbb{R}^n$; then $\Omega=P_n'\Omega P_n$ for every $\boldsymbol{P}_n\in\mathbb{P}_n$, i.e. Ω is invariant under $\boldsymbol{P}_n$ acting by congruence;

ii. Let $\mathcal{L}(\boldsymbol{Y})$ be left–exchangeable; then $\Omega=P_n'\Omega P_n$ for every $\boldsymbol{P}_n\in\mathbb{P}_n$;

iii. Let $\mathcal{L}(\boldsymbol{Y})$ be right–exchangeable; then $\Sigma=Q_k'\Sigma Q_k$ for every $\boldsymbol{Q}_k\in\mathbb{P}_k$.

Proof: Clearly $\mathcal{L}(\boldsymbol{y})=\mathcal{L}(\boldsymbol{P}_n'\boldsymbol{y})$ implies $\Omega=P_n'\Omega P_n$ for every $\boldsymbol{P}_n\in\mathbb{P}_n$, to give conclusion (i). Conclusions (ii) and (iii) follow as in Definition 1(iii), namely, $\mathrm{V}(P_n'\mathrm{Y}Q_k)=P_n'\Omega P_n\otimes Q_k'\Sigma Q_k$; and applying Conclusion (i) in succession to $\{\mathbf{Y}\to\boldsymbol{P}_n'\mathbf{Y}\}$ and $\{\mathbf{Y}\to\mathbf{Y}\boldsymbol{Q}_k\}$.

Classes of exchangeable errors

An early version having exchangeable rows on $\mathbb{F}_{n\times k}$ is $V(Y)=[\mathrm{I}_n\otimes(\Gamma-\Sigma)+J_n\otimes\Sigma;[(\Gamma-\Sigma),\Sigma]\in\mathbb{S}_k^+$, identified in[7] as an Exchangeable General Linear Model. This is a block–partitioned version of an equicorrelation matrix, but differing from matrices of type $V(\mathrm{Y})=\Omega\otimes\Sigma$ as considered here and listed in Table 1.

Table 1 Classes of matrices $\mathcal{H}_n^i(\Omega_i)$ for $\mathbf{Y}_i\in\mathbb{F}_{n\times k}$ as factors of $V(Y_i)=\Omega_i\otimes\Sigma$ having exchangeable rows, together with conditions $\Phi_i(\cdot)$ for Ω_i to be positive definite, where λ is of order $(n\times 1)$, $\tau_1=\lambda_1+\ldots+\lambda_n$ and $\tau_2=\Sigma_{i=1}^n(\lambda_i-\bar{\lambda})^2$

Class	Ω_i	$\Phi_i(\cdot)$	Source
$\mathcal{H}_n^1$	$\Omega(\lambda)=[I_n+1_n\lambda'+\lambda 1_n'-\bar{\lambda}J_n]$	$\{\tau_1>n\tau_2-1\}$	Jensen[8]
$\mathcal{H}_n^2$	$\Omega(\gamma,\lambda)=[\gamma I_n+1_n\lambda'+\lambda 1_n')]$	$\{\gamma>[(n\lambda'\lambda)^{\frac{1}{2}}-\tau_1]\}$	Baldessari[10]
$\mathcal{H}_n^3$	$\Omega(\rho)=[(1-\rho)I_n+\rho 1_n 1_n')]$	$\left\{\frac{-1}{(n+1)}\leq\rho\leq 1\right\}$	Halperin[9]

Remark 1: Given $\Omega(\gamma,\lambda)$, then $\Omega(\rho)=[(1-\rho)I_n+\rho J_n]$ follows on taking $\lambda=\theta 1_n$ and $\{\gamma=(1-\rho),2\theta=\rho\}$.

Essential properties may be summarized as follow.

Theorem 1: Consider the classes $\{\mathcal{C}=\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$ of Table 1, together with conditions Φ_i for $\mathbf{\Omega}_i$ to be positive definite. Then

i. The classes are closed under congruence by $\boldsymbol{P}_n$, i.e. for $\Omega_i\in H_n^1(\Omega_i)$, the matrices satisfy $\{P_n'\Omega_i P_n\in\mathcal{H}_n^i(\Omega_i)\}$, for each $\boldsymbol{P}_n\in\mathbb{P}_n$;

ii. For each $\{\mathcal{H}_n^i(\Omega_i)\}$, the conditions $\Phi_i(\cdot)$ that $\Omega_i\in\mathcal{H}_n^i(\Omega_i)$ be positive definite, are identical for all elements of the classes $\mathcal{C}$;

iii. Consider $\lambda=\lambda_0$ to be fixed as are $\tau_1^\dagger$ and $\tau_2^\dagger$. Corresponding to $\mathcal{H}_n^2(\Omega)$ is an equivalent subclass, namely $\mathcal{H}_2^\dagger$, as given by $\{\mathcal{H}_2^\dagger=[\gamma I_n+1_n\lambda_0'P_n'+P_n\lambda_0 1_n'];P_n\in\mathbb{P}_n\}$, having identical values for $\tau_1^\dagger$ and $\tau_2^\dagger$, consisting in number as $n!$ provided that elements of λ_0 are distinct.

Proof: (i) The conditions Φ_1 of Table 1 are from Theorem 2 of Jensen DR[8]; Φ_2 follows step–by–step on modifying that proof exclusive of $\bar{\lambda}$; and Φ_3 is given in Halperin M.[9] (ii) Closure properties for $\mathcal{H}_n^1$ and $\mathcal{H}_n^2$ follow with $A(\lambda)=1_n\lambda'+\lambda 1_n'$ since $[I_n+A(\lambda)]\to[I_n+A(\theta)]\in\mathcal{H}_n^1$ and $[\gamma I_n+A(\lambda)]\to[\gamma I_n+A(\theta)]\in\mathcal{H}_n^2$ with $\theta=P_n'\lambda$, and similarly $\mathcal{H}_n^3$ reproduces itself. Conclusion (iii) holds for $H_n^2(\Omega)$ since $\tau_1^\dagger$ and $\tau_2^\dagger$ are invariant under permutations of λ_0, and the members of $\mathcal{H}_2^\dagger$ clearly are generated from all $n!$ permutations of the elements of λ_0 if distinct.

Table 1 identifies additional parameters as required to achieve exchangeability. It is essential to examine the manner in which these may affect outcomes of the analysis, specifically, through the singular joint distribution of $[\hat{\mu},R]$ as functions of Y.

Repeated use is made of $V(A'Y)=A'\Omega A$ from Proposition 1(iii). In addition, $\mathbf{P}_{X_0}^\perp=\left[\boldsymbol{I}_n-\frac{1}{n}1_n 1_n'\right]$ is the idempotent projection operator onto the error span of the model $\mathcal{M}$.

Theorem 2: Given $\mathcal{L}(Y)\in\{N_{n\times k}(1_n\mu;\Omega\otimes\Sigma)\}$, consider $\mathcal{L}(\hat{\mu},R)$ under the classes $\mathcal{C}$ of Table 1. Then

i. $E[\hat{\mu},R]=[\mu,0]$ for each $\mathcal{H}_n^i\in\mathcal{C}$;

ii. The joint dispersion matrices $V(\hat{\mu},R)$ of order $(n+1)$ under the Table 1 classes are given respectively by

$$\overset{\Psi(\lambda)}{\begin{bmatrix}\frac{1}{n}+\bar{\lambda} & \lambda'P_{X_0}^\perp\\ P_{X_0}^\perp\lambda & P_{X_0}^\perp\end{bmatrix}};\overset{\Psi(\gamma,\lambda)}{\begin{bmatrix}\frac{\gamma}{n}+2\bar{\lambda} & \lambda'P_{X_0}^\perp\\ P_{X_0}^\perp\lambda & \gamma P_{X_0}^\perp\end{bmatrix}}\overset{\Psi(\rho)}{\begin{bmatrix}\frac{1+(n-1)\rho}{n} & 0\\ 0 & (1-\rho)P_{X_0}^\perp\end{bmatrix}} \quad (1)$$

Proof: Conclusion (i) follows from $E(\hat{\mu})=\frac{1}{n}1_n'E(Y)=\frac{1}{n}1_n'[1_n\mu]=\mu$, and $\mathrm{E}(\boldsymbol{R})=0$ by parallel arguments. Next let $L'=[\boldsymbol{L}_1',\boldsymbol{L}_2']=[\frac{1}{n}1_n',\mathbf{P}_{X_0}^\perp]$, so that $L'Y=[\hat{\mu},R]$ and $\{V(L'Y)=L'\Omega_i L\otimes\Sigma\}(**)$ with Ω_i as in Table 1. Substituting these in succession into expression $(**)$ gives the displayed matrices (1) for the classes $\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$, respectively.

In short, Theorem 2 catalogs the essentials of requiring that errors on $\mathbb{F}_{n\times k}$ be exchangeable as in Table 1. Both $(\hat{\mu},S)$ are affected in having properties discordant with those of the conventional $\mathcal{L}(Y)=N_{n\times k}(\mu,I_n\otimes\Sigma)$. Specifically, requiring that errors be exchangeable may serve to compromise the evidence contained in $\boldsymbol{S}$ with regard to Σ, to be examined subsequently. Details are collected in the following Table 2 as excerpted from Theorem 2.

Scale–invariance

This concept is central to establishing properties under independence as they may carry over to include exchangeable dependence. To these ends, associate with the classes $\mathcal{C}=\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$ the values $\kappa\in[1,\gamma,(1-\rho)]$ from the final row of Table 2.

Lemma 2: Let $\mathrm{T}(\mathbf{S})$ be scale invariant, i.e. $\mathrm{T}(\mathbf{S})=\mathrm{T}(\mathrm{c}\mathbf{S})$ for $\mathrm{c}\neq 0$; and consider these as they may apply in the exchangeable classes $\mathcal{C}=\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$ of Table 1.

i. The scale parameters of $\mathcal{L}(\nu\mathbf{S})$ are respectively $\{\kappa\mathbf{\Sigma};\kappa\in[1,\gamma,(1-\rho)\}$ for the classes of Table 1;

ii. Properties of $\mathrm{T}(\mathbf{S})$ are identical to those for $\{\mathcal{L}(Y)\in\{N_{n\times k}(1_n\mu;I_n\otimes\Sigma)\}$, for each of the exchangeable classes $\mathcal{C}$.

Proof: Conclusion (i) is from Table 2 as noted. The proof for (ii) hinges on scaling properties of Wishart matrices, namely, that $\mathbf{R}'\mathbf{R}=\nu\mathbf{S}$, so that if $\mathcal{L}(\nu S)=W_k(\nu,\kappa\Sigma,0)$ as in Table 2, then $\mathcal{L}(\nu S/\kappa)=W_k(\nu,\Sigma,0)$, the default state. Accordingly, infer that $\mathrm{T}(\mathbf{S})$ behaves as if from $\mathcal{L}(\nu S)=W_k(\nu,\kappa\Sigma,0)$ in the third row of Table 2, and $\mathrm{T}(\mathbf{S}/\kappa)$ behaves as if from $\mathcal{L}(\nu S/\kappa)=W_k(\nu,\Sigma,0)$. But T is scale–invariant, so that $\mathrm{T}(\mathbf{S})=\mathrm{T}(\mathbf{S}/\kappa)$, as if from $\{\mathcal{L}(Y)\in\{N_{n\times k}(1_n\mu;I_n\otimes\Sigma)\}$ to complete a proof.

Tests for μ

A complement to estimation is hypothesis testing under exchangeable errors. First consider μ. For $V(Y)=I_n\otimes\Sigma$, recall that

i. $V(\hat{\mu})=\frac{1}{n}\Sigma$;

ii. $[\hat{\mu},S]$ are mutually independent; and

iii. Hotelling's[11] test for $\mathrm{H}_0:\mu=\mu_0$ vs $\mathrm{H}_1:\mu\neq\mu_0$ utilizes the statistic

$$T^2=n(\hat{\mu}-\mu_0)S^{-1}(\hat{\mu}-\mu_0)' \tag{2}$$

with distribution $\mathcal{L}(T^2)=T_k^2(\nu,\theta)$ of order k having $\nu=(\mathrm{n}-1)$ degrees of freedom and noncentrality parameter θ. Under the error classes of Table 1, the principal negative finding of this study is the following.

Lemma 3: Consider $\hat{\mu}$ in the classes $\mathcal{C}=\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$, together with T^2 for testing $\mathrm{H}_0:\mu=\mu_0$ vs $\mathrm{H}_0:\mu=\mu_0$.

i. That $(\hat{\mu},S)$ are independent is met only in the class $\mathcal{H}_n^3$;

ii. Replacing n in Equation (2) are reciprocals of $\left(\frac{1}{n}+\bar{\lambda}\right),\left(\frac{\gamma}{n}+2\bar{\lambda}\right),\left(\frac{1+(n-1)\rho}{n}\right)$, and these typically are unknown;

iii. In short, the classical tests for $\hat{\mu}$ are unsupported in the exchangeable error classes $\mathcal{C}$.

Proof: (i) The independence of $(\hat{\mu},S)$, namely $Cov(\hat{\mu},R)=0$, is met only in the class $\mathcal{H}_n^3$ in Theorem 2, unless $\lambda\in S_{pn}(X_0)$ for both $\mathcal{H}_n^1$ and $\mathcal{H}_n^2$ in Equation (1), in which case $\lambda' P_{X_0}^{\perp}=0$. Conclusion (ii) follows from Theorem 2 and Table 2, and Conclusion (iii) follows in summary.

Table 2 Properties of $\boldsymbol{R}=\mathbf{P}_{X_0}^{\perp}\mathbf{Y}$ and $\nu\boldsymbol{S}=\boldsymbol{R}'\boldsymbol{R}$, where $\mathbf{P}_{X_0}^{\perp}=\left[\boldsymbol{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right]$; moreover, the distribution $W_k(\nu,\Sigma,0)$ is central Wishart of order k, having $\nu=(n-1)$ degrees of freedom and scale parameters Σ

Item	$\mathcal{H}_n^1$	$\mathcal{H}_n^2$	$\mathcal{H}_n^3$
$\mathcal{L}(\hat{\mu})$	$N_{1\times k}(\mu,(\frac{1}{n}+\bar{\lambda})\Sigma)$	$N_{1\times k}(\mu,(\frac{\gamma}{n}+2\bar{\lambda})\Sigma)$	$N_{1\times k}(\mu,(\frac{1+(n-1)}{n}\rho)$
$\mathcal{L}(R)$	$N_{n\times k}(0,P_{X_0}^{\perp}\otimes\Sigma)$	$N_{n\times k}(0,P_{X_0}^{\perp}\otimes\gamma\Sigma)$	$N_{n\times k}(0,P_{X_0}^{\perp}\otimes(1-\rho)\Sigma)$
$\mathcal{L}(\nu\boldsymbol{S})$	$W_k(\nu,\Sigma,0)$	$W_k(\nu,\gamma\Sigma,0)$	$W_k(\nu,(1-\rho),\Sigma,0)$
$E(\boldsymbol{S})$	Σ	$\gamma\Sigma$	$(1-\rho)\Sigma$

Inferences for Σ

Estimation

The dispersion matrix $\{V(Y_i)=\omega_{ii}\Sigma\}$ within the rows of $\mathbf{Y}$, and the cross–covariances $\{Cov(Y_i,Y_j)=\omega_{ij}\Sigma\}$ between rows, all depend on Σ. In addition to properties of $\mathbf{S}=\mathbf{R}'\mathbf{R}/(n-1)$ as reported in Theorem 2 and Table 2, let $[\mathrm{s}_1,\mathrm{s}_2,\mathrm{s}_3]$ be the error mean squares for the classes $\mathcal{C}=\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$. Essential features are that $\{\mathcal{L}(\nu,S_i)=W_k(\nu,\kappa_i\Sigma,0)\}$ for $\nu=(\mathrm{n}-1)$ and $\kappa_i\in[1,\gamma,(1-\rho)]$ for the classes $\mathcal{C}$. Thus $\mathbf{S}_1$ is unbiased for Σ, whereas $(\mathrm{s}_2,\mathrm{s}_3)$ are biased by the factors $[\gamma,(1-\rho)]$. Moreover, as measures of scatter, the generalized variances are related as $|S_2|=\gamma^k|S_1|$ and $|S_3|=(1-\rho)^k|S_1|$, whereas the condition numbers $\{\mathrm{Cnd}(\mathbf{S}_i);\mathrm{i}=1,2,3\}$ are identical.

Hypothesis tests

Five tests, historically devised and subsequently used under $\{\mathcal{L}(Y)\in\left\{N_{n\times k}\left(1_n\mu;I_n\otimes\Sigma\right)\right\}$ are listed in Table 3, to include statements of hypotheses, commonly used test statistics, and references.

As to exchangeable dependence, it remains to identify those of Table 3 that remain viable in the exchangeable classes of Table 1.

Theorem 3: Consider the tests for Σ as in Table 3 for the classes $\mathcal{C}=\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$ of Table 1, in lieu of the conventional $\{\mathcal{L}(Y)\in\left\{N_{n\times k}\left(1_n\mu,I_n\otimes\Sigma\right)\right\}$.

i. All statistics of Table 3 are scale–invariant;

ii. For the classes $\mathcal{C}$, properties of the tests of Table 3 are identical to those for $\{\mathcal{L}(Y)\in\left\{N_{n\times k}\left(1_n\mu,I_n\otimes\Sigma\right)\right\}$, independently of $\kappa\in[1,\gamma,(1-\rho)]$.

Proof: As before, $\kappa\in[1,\gamma,(1-\rho)]$ are the scale parameters for $\mathbf{S}$ in $\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$. Conclusion (i) is apparent, where for $H_5:S_0=S\Sigma_0^{-1}$, we find on rescaling $\mathrm{Y}\to\kappa\mathrm{Y}$ that $\boldsymbol{S}\to\kappa^2\boldsymbol{S}$ and $\Sigma_0\to\kappa^2\Sigma_0$, leaving $\boldsymbol{S}_0$ to be scale–invariant. Conclusion (ii) follows on applying conclusion (i) in order to verify the scale–invariance and applicability of Lemma 2.

Remark 2: These tests accordingly exhibit genuinely nonparametric features, in that each applies for structured distributions in the classes $\{\mathcal{H}_n^1,\mathcal{H}_n^2,\mathcal{H}_n^3\}$ beyond that of the conventional $\{\mathcal{L}(Y)\in\left\{N_{n\times k}\left(1_n\mu,I_n\otimes\Sigma\right)\right\}$.

Exact distributions of the Table 3 statistics $\{u_i\}$ rarely are known, supported instead by approximations, namely, $\{u_i\to u_i'=c_i\phi(u_i);\phi\in[u_i,\ln u_i]\}$, such that $\mathcal{L}(u_i')\approx\chi^2_{\nu_i}$, namely, approximately chi–squared having ν_i degrees of freedom. Details are found in Sections 7.2.1 and 7.2.2 of Rencher[12] and 7.3 of Morrison.[13] These details are omitted here in the interests of brevity, but suffice to say, those approximations all apply in the exchangeable error classes of this study.

Table 3 Selected hypotheses regarding Σ; commonly used test statistics; references R to Rencher[12] and M to Morrison[13]

Item	$H_0 : \Sigma =$	Test statistic (u_i)	Reference
H_1	$\sigma^2 \boldsymbol{I}_k$	$\left[\dfrac{\lvert \boldsymbol{S} \rvert}{(tr\, \boldsymbol{S}/k)^k}\right]$	R¶7.2.2
H_2	$\sigma^2[(1-\rho)\boldsymbol{I}_k + \rho \boldsymbol{J}_k]$	$\left[\dfrac{\lvert \boldsymbol{S} \rvert}{\lvert s^2[(1-r)\boldsymbol{I}_k + r\boldsymbol{J}_k]\rvert}\right]$	R¶7.2.3
H_3	$[I_k + 1_k \lambda' + \lambda 1_k' - \bar{\lambda} J_k]$	$\dfrac{(\kappa)^\kappa \lvert \boldsymbol{C'SC} \rvert}{(tr\boldsymbol{C'SC})^\kappa}$	M¶7.3
H_4	$Diag(\Sigma_{11},\ldots,\Sigma_{rr})$	$\dfrac{\lvert \boldsymbol{S} \rvert}{\lvert \boldsymbol{S}_{11} \rvert \lvert \boldsymbol{S}_{22} \rvert \cdots \lvert \boldsymbol{S}_{rr} \rvert}$.	R¶7.4.2
H_5	Σ_0	$v[-\ln \lvert S\Sigma_0^{-1} \rvert + tr(S\Sigma_0^{-1}) - k]$	R¶7.2.1

Legend $\kappa = (k-1)$; $S = [s_{ij}] \rightarrow \{s^2 = \frac{1}{k}\sum_{i=1}^{k}; r = \left[\frac{1}{k(k-1)}\Sigma_{i \neq j} s_{ij}\right] / s^2$ and $C'[\kappa \times k]$ consists of κ linear contrasts

Correlation analyses

Here sample entities depend on $\boldsymbol{S} = [s_{ij}]$, corresponding parameters are identical functions of Σ. To these ends take $\boldsymbol{Y} = [\boldsymbol{Y}_1, \boldsymbol{Y}_2]$ of orders $\{(n \times s), (n \times t); s \leq t\}$, and partition $\boldsymbol{S}(k \times k)$ as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{I}_s & \mathbf{G} \\ \mathbf{G}' & \mathbf{I}_t \end{bmatrix}; \mathbf{G} = \mathbf{S}_{11}^{-\frac{1}{2}} \mathbf{S}_{12} \mathbf{S}_{22}^{-\frac{1}{2}} (s \times t). \quad (3)$$

Then $\left\{ r_{ij} = s_{ij} \Big/ s_{ii}^{\frac{1}{2}} s_{jj}^{\frac{1}{2}} \right\}$ are simple correlations; the singular values $\sigma(G)$ are the canonical correlations $\varrho = [\varrho_1, \varrho_2, \ldots \varrho_s]$ and the multiple correlations are defined at $s=1$. Again note that these were derived historically and subsequently used under the independence model $\{\mathcal{L}(Y) \in \{N_{n \times k}(1_n \mu, I_n \otimes \Sigma)\}$. The question again arises as to whether exchangeable errors may have compromised correlative evidence in $\boldsymbol{S}$ regarding Σ Results to the contrary are the substance of the following.

Theorem 4: Given $\mathcal{L}(Y)$ in the exchangeable classes $\mathcal{C} = \{\mathcal{H}_n^1, \mathcal{H}_n^2, \mathcal{H}_n^3\}$; consider effects on correlation analyses as prescribed under $\mathcal{L}(Y^*) = N_{n \times k}(1_n \mu, I_n \otimes \Sigma)$.

i. Then for all $\mathcal{L}(\boldsymbol{Y}) \in \mathcal{C}$, the entities $\{r_{ij}\}$ and their properties are identical to those for $\mathcal{L}(\boldsymbol{Y}^*)$;

ii. In like manner, for all $\mathcal{L}(\boldsymbol{Y}) \in \mathcal{C}$, properties of multiple and canonical correlations are identical to those for $\mathcal{L}(Y^*)$;

iii. In short, conventional correlation analyses are preserved despite requiring that errors be exchangeable in $\mathcal{C}$.

Proof: The claims again rest on the fact that sample correlations are scale–invariant functions of $\mathbf{Y}$ and $\mathbf{S}$. Conclusions (i), (ii) and (iii) now follow from Lemma 2.

Factor analyses (FA's)

Within the scope of psychometric, sociometric, and humanistic endeavors, the *FA* paradigm postulates that $\Sigma = \Lambda'\Lambda + \Psi$ such that elements of $\{\Lambda(s \times k); s < k\}$ comprise the factor loadings, and $\Psi = Diag(\psi_1, \ldots, \psi_k)$ the unique variances. In particular, the diagonal elements of Σ are $\{\sigma_{ii} = h_i^2 + \psi_i; i = 1, \ldots, k\}$ where $\{h_i^2 = \lambda_{i1}^2 + \lambda_{i2}^2 + \ldots + \lambda_{is}^2; i = 1, \ldots, k\}$ are the communalities. The analysis begins with $S = \widehat{\Lambda}'\hat{\Lambda} + \Psi$, typically utilizing maximum likelihood estimation as in Chapter 13 of Rencher.[12] An initial solution $\hat{\Lambda}$ eventually is rotated so as to achieve further desirable properties, since the loadings Λ are non–unique.

For the case that $\{\mathcal{L}(Y) \in \{N_{n \times k}(1_n \mu, I_n \otimes \Sigma)\}$, the normal–theory likelihood ratio for testing $H_0 : \Sigma = \Lambda'\Lambda + \Psi$ vs $H_1 : \Sigma \neq \Lambda'\Lambda + \Psi$ is

$$\left[n - \frac{2k + 4s + 11}{6}\right] \ln \left[\frac{\lvert \widehat{\Lambda}'\hat{\Lambda} + \hat{\Psi} \rvert}{\lvert S \rvert}\right] \quad (4)$$

and referred to upper critical values of the approximating distribution, namely, χ_v^2 with $v = \left[(k-s)^2 - k - s/2\right]$ as in expression (13.47) of Rencher.[12] These were derived historically and used subsequently for the case that $\{\mathcal{L}(Y) \in \{N_{n \times k}(1_n \mu, I_n \otimes \Sigma)\}$.

The extent to which the foregoing algorithm may be applied more generally, to encompass exchangeable errors, is examined in the following.

Theorem 5: Consider the statistic (4) for testing the *FA* model in the classes $\mathcal{C} = \{\mathcal{H}_n^1, \mathcal{H}_n^2, \mathcal{H}_n^3\}$, as developed and prescribed for $\mathcal{L}(Y^*) = N_{n \times k}(1_n \mu, I_n \otimes \Sigma)$. Then

i. For each distribution $\mathcal{L}(\boldsymbol{Y}) \in \mathcal{C}$, properties of tests using (4) are identical to those under $\mathcal{L}(\boldsymbol{Y}^*)$.

Proof: As the statistic (4) is scale–invariant, the conclusion again follows from Lemma 2.

Conclusion

In retrospect, taking the conventional $V(Y) = I_n \otimes \Sigma$ remains an enduring artefact of statistical practice. Exchangeable dependence, where $V(Y) = \Omega \otimes \Sigma$, is a radical departure, albeit on occasion as being itself fundamental to correct statistical practice. Foundations trace to Johnson[3]; extensions encompass matrices in $\mathbb{F}_{n\times k}$ and stochastic sequences in various domains. Representations for two–way arrays include (i) functions of *iid* scalars as in Aldous DJ[6] and the related studies[14,15]; and (ii) as limits of finite exchangeable sequences as in Ivanoff BG.[16] Ivanoff BG[16] for rectangular arrays. Marshall & Olkin[17] demonstrated that Schur–concave joint density functions on $\mathbb{R}^n$ are exchangeable; Shaked & Tong[18] superimposed partial orderings on exchangeable arrays; and Seneta[19] sought to approximate joint probabilities of equicorrelated vectors in $\mathbb{R}^n$ in terms of marginal probabilities and the correlation parameter ρ Functional limit theorems for row and column arrays were studied in Ivanoff BG.[16] Kallenberg[20] examined ergodic properties of exchangeable arrays generated as multivariate samples from a stationary process. In reliability studies, an exchangeable array is considered in Spizzichino F, et al.[21] as deriving from a hierarchical model having multivariate negative aging. In addition, a multivariate lognormal frailty model for exchangeable failure time data, having marginal Weibull lifetime distributions, is considered in Stefanescu C.[22]

Alternative to our studies is equation (1) of Arnold[7] having the linear structure of our model $\mathcal{M}$ but differing in dispersion. Arnold's approach differs in reducing his model to a canonical form. Nonetheless, Arnold's assessment of $\hat{\mu}$ serves to confirm our findings in Lemma 3. On the other hand, our examination of Σ, its sample version $\boldsymbol{S}$, and other second–moment properties, find no parallel in Arnold's studies. In continuation of those studies, Roy & Fonseca[23] sought to extend equation (1), considered as a two–level array, to encompass three levels.

Antecedents to the present study include $\Omega(\gamma,\lambda)$ in Table 1 from Baldessari[10] in lieu of $\sigma^2\mathbf{I}_n$ in the Analysis of Variance; and characterized in[24–26] as the class of all within-subject dispersion matrices preserving the validity of conventional F–tests in the analysis of repeated measurements. Moreover, structured matrices of an earlier vintage include the Euclidean distance matrices of Gower,[27] namely $D(\lambda) = \left[D + 1_n\lambda' + \lambda 1_n'\right]$, with D diagonal, having applications to linear inference as found in Farebrother.[28]

In summary, our studies have sought to cover a diversity of topics in multivariate statistical inference from a further perspective, namely, that of exchangeable errors. But at the same time, to acknowledge and to pursue the prospects that requiring exchangeability may serve to compromised the meanings attributed to sample evidence. Specifically, references abound for the vast array of multivariate normal procedures described here as classical, including those amenable to selected exchangeable distributions as shown here. Of the many topics not covered, interested readers are encouraged to undertake further investigations using and adding to the analytical principles demonstrated here.

Acknowledgements

None.

Conflicts of interest

The author declare that there is no conflicts of interest.

Funding

None.

References

1. Jensen DR, Ramirez DE. Anomalies in the analysis of calibrated data. *J Statist Computat Simul*. 2009;79:299–314.
2. Jensen DR, Ramirez DE. Irregularities in X(Y) from Y(X) in linear regression. *J Statist Computat Simul*. 2012;83(10):1807–1828.
3. Johnson W. Logic, Part III: The Logical Foundations of Science. Cambridge UK: Cambridge University Press; 1924.
4. Jensen DR, Good IJ. Invariant distributions associated with matrix laws under structural symmetry. *J Royal Statist Soc B*. 1981;43(3):327–332.
5. Finetti B. Forecasting: its logical laws, its subjective sources. *Annals of the Henri Poincaré Institute*. 1937;7(1):1–68.
6. Aldous DJ. Representations for partially exchangeable arrays of random variables. *J Multivariate Anal*. 1981;11(4):581–598.
7. Arnold S. Linear models with exchangeably distributed errors. *J Amer Statist Assoc*. 1979;74(365):194–199.
8. Jensen DR. Structured dispersion and validity in linear inference. *Linear Algebra Applic*. 1996;249:189–196.
9. Halperin M. Normal regression theory in the presence of intra-class correlation. *Ann Math Statist*. 1951;22:573–580.
10. Baldessari B. Analysis of variance of dependent data. *Statistica (Bologna)*. 1966;26:895–903.
11. Hotelling H. The generalization of Student's ratio. *Ann Math Statist*. 1931;2:360–378.
12. Rencher A. *Methods of Multivariate Analysis*. 2nd ed. New York: Wiley; 2002.
13. Morrison DF. *Multivariate Statistical Methods*. 3rd ed. New York: McGraw–Hill; 1990.
14. Kallenberg O. On the representation theorem for exchangeable arrays. *J Multivariate Anal*. 1989;30:137–154.
15. Lynch J. Canonical row–column–exchangeable arrays. *J Multivariate Anal*. 1984;15:135–140.
16. Ivanoff BG, Weber NC. Functional limit theorems for row and column exchangeable arrays. *J Multivariate Anal*. 1995;55:133–148.
17. Marshall AW, Olkin I. Majorization in multivariate distributions. *Ann Statist*. 1974;2:1189–1200.
18. Shaked M, Tong YL. Some partial orderings of exchangeable random variables by positive dependence. *J Multivariate Anal*. 1985;17:333–349.
19. Seneta E. Multivariate probability in terms of marginal probability and correlation coefficient. *Biometrical J*. 1987;29:375–380.
20. Kallenberg O. Multivariate sampling and the estimation problem for exchangeable arrays. *J Theor Probab*. 1999;12:859–883.
21. Spizzichino F, Torrisi G. Multivariate negative aging in an exchangeable model of heterogeneity. *Statist Probab Letters*. 2001;55:71–82.
22. Stefanescu C, Turnbull B. Multivariate frailty models for exchangeable survival data with covariates. *Technometrics*. 2006;48:411–417.

23. Roy A, Fonseca M. Linear models with doubly exchangeable distributed errors. *Commun. Statist: Theory and Methods*. 2012;41;2545–2569.
24. Bock R. Multivariate analysis of variance of repeated measurements. In: Harris CW, editor. *Problems of Measuring Change*. Madison: Wisconsin: University of Wisconsin Press; 1963:85–103.
25. Huynh H, Feldt L. Conditions under which mean square ratios in repeated measurements designs have exact F–distributions. *J Amer Statist Assoc*. 1970;65:1582–1589.
26. Rouanet H, Lépine D. Comparison between treatments in a repeated–measurement design: ANOVA and multivariate methods, British *J Math Statist Psychol*. 1970;23:147–163.
27. Gower J. Euclidean distance geometry. *Math Sci*. 1982;7:1–14.
28. Farebrother R. Linear unbiased approximators of the disturbances in the standard linear model. *Linear Algebra Applic*. 1985;67:259–274.