

Accepted Manuscript

Hotelling's T^2 tests in paired and independent survey samples: An efficiency comparison

Ludwig Baringhaus, Daniel Gaigall

PII: S0047-259X(16)30134-8

DOI: <http://dx.doi.org/10.1016/j.jmva.2016.11.004>

Reference: YJMVA 4185

To appear in: *Journal of Multivariate Analysis*

Received date: 12 May 2016



Please cite this article as: L. Baringhaus, D. Gaigall, Hotelling's T^2 tests in paired and independent survey samples: An efficiency comparison, *Journal of Multivariate Analysis* (2016), <http://dx.doi.org/10.1016/j.jmva.2016.11.004>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Hotelling's T^2 tests in paired and independent survey samples: An efficiency comparison

Ludwig Baringhaus, Daniel Gaigall

*Institut für Mathematische Stochastik, Leibniz Universität Hannover,
Postfach 60 09, D-30060 Hannover, Germany*

Abstract

Hotelling's T^2 tests in paired and independent survey samples are compared using the traditional asymptotic efficiency concepts of Hodges–Lehmann, Bahadur and Pitman, as well as through criteria based on the volumes of corresponding confidence regions. Conditions characterizing the superiority of a procedure are given in terms of population canonical correlation type coefficients. Statistical tests for checking these conditions are developed. Test statistics based on the eigenvalues of a symmetrized sample cross-covariance matrix are suggested, as well as test statistics based on sample canonical correlation type coefficients.

Key words: Hotelling's T^2 test, relative efficiency, volumes of confidence sets, sample covariance matrix, canonical correlation type coefficients, empirical eigenvalues

2000 MSC: 62H15, 62K05

1. Introduction

Let $n, d \in \mathbb{N}$, $n > d$. Consider independent and identically distributed pairs of random vectors $(X_1, Y_1), \dots, (X_{2n}, Y_{2n})$ with values in $\mathbb{R}^d \times \mathbb{R}^d$ and distribution $\mathcal{L}(X_1, Y_1)$. In many situations, one can choose between the experimental design of a paired survey sample with observations of the independent and identically distributed random vectors

$$(X_1, Y_1), \dots, (X_n, Y_n) \tag{1.1}$$

and the experimental design of independent survey samples with observations of the two independent blocks of independent and identically distributed random vectors

$$X_1, \dots, X_n \quad \text{and} \quad Y_{n+1}, \dots, Y_{2n}.$$

Email addresses: lbaring@stochastik.uni-hannover.de (Ludwig Baringhaus),
gaigall@stochastik.uni-hannover.de (Daniel Gaigall)

Preprint submitted to Journal of multivariate analysis

November 1, 2016

In the latter case, one can also set $X'_i = X_i$ and $Y'_i = Y_{n+i}$, $i = 1, \dots, n$, so that one then deals with observations of the independent and identically distributed random vectors

$$(X'_1, Y'_1), \dots, (X'_n, Y'_n). \quad (1.2)$$

Thus, formally, there is a sampling scheme involved here, just as in the paired sample case. Denoting by $\mu \otimes \nu$ the product of probability measures μ and ν , one can set

$$\nu_X = \mathcal{L}(X_1) = \mathcal{L}(X'_1) \quad \text{and} \quad \nu_Y = \mathcal{L}(Y_1) = \mathcal{L}(Y'_1),$$

and hence $\mathcal{L}(X'_1, Y'_1) = \mathcal{L}(X'_1) \otimes \mathcal{L}(Y'_1)$, while in the paired sample case, one generally has $\mathcal{L}(X_1, Y_1) \neq \mathcal{L}(X_1) \otimes \mathcal{L}(Y_1)$.

As both sampling schemes involve the same number n of observations, it is natural to ask which of the two designs is best. The purpose of this paper is to address this issue in the special case where the joint distribution $\mathcal{L}(X_1, Y_1)$ is multivariate normal and where, based on each of the sampling schemes (1.1) and (1.2), one wishes to test the following hypotheses:

$$\mathcal{H} : \nu_X = \nu_Y, \quad \mathcal{K} : \nu_X \neq \nu_Y. \quad (1.3)$$

To fix ideas and provide motivation, let us first deal with a very special bivariate normal case.

Example 1.1. Assume $d = 1$ and

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathcal{N}_2 \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \rho \\ \sigma^2 \rho & \sigma^2 \end{pmatrix} \right],$$

where $\sigma^2 > 0$ and $\rho \in (-1, 1)$ are known and $a, b \in \mathbb{R}$ are unknown. The law of the independent samples is then given by

$$\begin{pmatrix} X'_1 \\ Y'_1 \end{pmatrix} \sim \mathcal{N}_2 \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right].$$

Setting $\mu = a - b$, one can then write the testing problem (1.3) as $\mathcal{H} : \mu = 0$, $\mathcal{K} : \mu \neq 0$. Suitable test statistics are

$$T_n = \frac{\sqrt{n} |\bar{X}_n - \bar{Y}_n|}{\sqrt{2\sigma^2(1-\rho)}} \quad \text{and} \quad T'_n = \frac{\sqrt{n} |\bar{X}'_n - \bar{Y}'_n|}{\sqrt{2\sigma^2}}.$$

For given $\alpha \in (0, 1)$, denote by $u_{1-\alpha/2}$ the quantile of order $1 - \alpha/2$ of the standard normal distribution, denoted $\mathcal{N}(0, 1)$. The statistical tests with rejection regions $\{T_n > u_{1-\alpha/2}\}$ and $\{T'_n > u_{1-\alpha/2}\}$ are of exact size $\alpha \in (0, 1)$. We compare the tests by means of the corresponding error probabilities of the second kind $\beta_{T,n}$ and $\beta_{T',n}$. Defining

$$\psi_y(x) = \Phi(y - x) - \Phi(-y - x), \quad x \in \mathbb{R}, \quad y \in (0, \infty),$$

where Φ is the distribution function of $\mathcal{N}(0, 1)$, we have, for given $\mu \neq 0$,

$$\beta_{T,n} = \psi_{u_{1-\alpha/2}} \left\{ \frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}} \mu \right\} \quad \text{and} \quad \beta_{T',n} = \psi_{u_{1-\alpha/2}} \left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}} \mu \right).$$

It is easily seen that

$$\beta_{T,n} \begin{cases} < \\ = \\ > \end{cases} \beta_{T',n} \text{ for each } n \in \mathbb{N} \text{ if and only if } \rho \begin{cases} > \\ = \\ < \end{cases} 0.$$

Thus the paired survey sample experimental design is preferable if X_1 and Y_1 are positively correlated. If the correlation is negative, the independent sampling scheme dominates. The quality of both procedures is equal if the correlation is zero. Comparing the length $2n^{-1/2}u_{1-\alpha/2}\sqrt{2\sigma^2(1-\rho)}$ of the related confidence interval for μ in paired samples with the length $2n^{-1/2}u_{1-\alpha/2}\sqrt{2\sigma^2}$ of that in independent samples, the same conclusion is drawn.

In general, however, crisp and blanket conclusions such as these on the superiority of one of the sampling schemes over the other are not possible. There is only sparse literature dealing with this problem on a rigorous basis. For 2×2 contingency tables, comparisons of McNemar's one-sided symmetry test (paired samples) and Fisher's one-sided two-sample exact test (independent samples) were conducted by Knüsel [11] using Pitman's asymptotic efficiency concept and Frischmuth [7] using the Hodges–Lehmann efficiency concept. The latter author presents also a result obtained from the comparison of the one-sided one-sample t -test (paired samples) and the one-sided two-sample t -test (independent samples) by means of the asymptotic Hodges–Lehmann efficiency; see also Section 4. Establishing properties of the power functions of these t -tests for finite sample sizes, Pollak and Cohen [21] came to divided conclusions regarding the superiority of either one of these t -tests.

Before dealing with Hotelling's T^2 tests in paired and independent survey samples, it is helpful to state some useful facts; this is done in Section 2. The basis of our comparisons are the asymptotic efficiency concepts of Hodges–Lehmann, Bahadur and Pitman, as well as criteria based on volumes of confidence sets. Details on these notions are given in Section 3. An efficiency comparison of the Hotelling tests is presented in Section 4. Conditions characterizing the superiority of a procedure are derived in terms of population canonical correlation type coefficients. Statistical tests checking these conditions are developed and discussed in Section 5. Test statistics based on the eigenvalues of $\{S_n^{12} + (S_n^{12})^\top\}/2$, where S_n^{12} is the sample cross-covariance matrix of the random vectors $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ as well as tests based on sample canonical correlation type coefficients, i.e., the eigenvalues of $\{S_n^{12} + (S_n^{12})^\top\}(S_n^{11} + S_n^{22})^{-1}$ with S_n^{11} and S_n^{22} as the sample covariance matrix of $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$, respectively, are suggested. Section 6 contains a numerical study on the power performance of the two Hotelling tests in the finite-sample case; these findings are in line with our theoretical considerations. The result of a small simulation study on the power performance of one of the superiority tests is also presented there.

2. Hotelling's T^2 tests

For any $k \in \mathbb{N}$, denote by

$$\mathcal{P}_k = \{A \in \mathbb{R}^{k \times k} : A \text{ is symmetric positive definite}\}$$

the set of symmetric positive definite $k \times k$ matrices. In what follows, we assume that

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathcal{N}_{2d} \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix} \right], \quad (2.1)$$

where $a, b \in \mathbb{R}^d$ are unknown and $\Sigma, \Pi \in \mathbb{R}^{d \times d}$, chosen in such a way that

$$\Xi = \begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix} \in \mathcal{P}_{2d}, \quad (2.2)$$

are also unknown. Note that in this case the common covariance matrix Σ of X_1 and Y_1 is also positive definite. Obviously,

$$\begin{pmatrix} X'_1 \\ Y'_1 \end{pmatrix} \sim \mathcal{N}_{2d} \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} \right].$$

With $\mu = a - b$, the testing problem (1.3) can be written as

$$\mathcal{H} : \mu = 0, \quad \mathcal{K} : \mu \neq 0. \quad (2.3)$$

We consider testing (2.3) by using Hotelling T^2 tests in paired and independent survey samples. A detailed description is given below. However, let us first give a preview of the main findings obtained with the efficiency comparisons of these two tests. Set $\Lambda = 2\Sigma - (\Pi + \Pi^\top) \in \mathcal{P}_d$. It turns out that the efficiencies of Hotelling's T^2 tests in independent samples with respect to Hotelling's T^2 test in paired samples can essentially be described by the ratio of two quadratic forms, viz.

$$\eta^\top (2\Sigma)^{-1} \eta / \eta^\top \Lambda^{-1} \eta,$$

ranging between 0 and 2, where the parameter $\eta \in \mathbb{R}^d$ is determined by the alternative used for the calculation of the efficiencies, or are given by the ratio of the related determinants,

$$\det(2\Sigma)^{-1} / \det \Lambda^{-1},$$

ranging between 0 and 2^d . Moreover, it will be shown that the superiority of one of the sampling designs can be characterized by the eigenvalues $\lambda_1(\Pi, \Sigma) \leq \dots \leq \lambda_d(\Pi, \Sigma)$ of the symmetric matrix $\Sigma^{-1/2}(\Pi + \Pi^\top)\Sigma^{-1/2}/2$. See the discussion in Section 4 for details.

Note that in the case $\Pi = \Pi^\top$, $|\lambda_1(\Pi, \Sigma)|, \dots, |\lambda_d(\Pi, \Sigma)|$ are the population canonical correlation coefficients of the partitioned matrix Ξ . Therefore, we are led to speak of the canonical correlation type coefficients $\lambda_1(\Pi, \Sigma), \dots, \lambda_d(\Pi, \Sigma)$. In the special case $d = 1$,

$$\eta^\top (2\Sigma)^{-1} \eta / \eta^\top \Lambda^{-1} \eta = \det(2\Sigma)^{-1} / \det \Lambda^{-1} = 1 - \rho,$$

where ρ is the correlation coefficient of X_1 and Y_1 . Thus, as in Example 1.1, paired survey sampling turns out to be the better experimental design if the correlation between X_1 and Y_1 is positive. If the correlation is negative, the independent survey sampling dominates. The quality of both procedures is equal if the correlation is zero. See also Frischmuth [7], who compared the one-sided one-sample t -test (paired samples) and the one-sided two-sample t -test (independent samples) by means of the asymptotic Hodges–Lehmann efficiency.

Let us now come back to a short description of the Hotelling tests. Under the sampling scheme (1.1) it is convenient to use Hotelling's T^2 test statistic based on the differences $Z_i = X_i - Y_i, i = 1, \dots, n$. Of course, $Z_1 \sim \mathcal{N}_d(\mu, \Lambda)$. Hotelling's T^2 test statistic (modified so as to have the F -distribution) is

$$T_n = \frac{n(n-d)}{d(n-1)} \bar{Z}_n^\top (S_n^Z)^{-1} \bar{Z}_n,$$

where

$$S_n^Z = \frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z}_n)(Z_i - \bar{Z}_n)^\top$$

is the sample covariance matrix computed from $Z_1, \dots, Z_n$. Since $n > d$, S_n^Z is positive definite with probability 1. The distribution of T_n is the same as that of

$$\frac{W_1/d}{W_2/(n-d)} \sim F_{d,n-d}(\delta),$$

where $\delta = n\mu'\Lambda^{-1}\mu$ and $W_1 \sim \chi_d^2(\delta)$, $W_2 \sim \chi_{n-d}^2$ are independent; see Muirhead [19], Theorems 1.3.6 and 6.3.1. For given $\alpha \in (0, 1)$, denote by $F_{d,n-d}^{1-\alpha}$ the quantile of order $1 - \alpha$ of the (central) F -distribution with d and $n - d$ degrees of freedom. The rejection region of Hotelling's T^2 test is

$$\{T_n \geq F_{d,n-d}^{1-\alpha}\}. \quad (2.4)$$

It is of exact size α , and it is consistent. Based on the sample $Z_1, \dots, Z_n$, it is the likelihood ratio test for the testing problem (2.3); see Muirhead [19], Theorem 6.3.2. Define

$$\text{GL}(d, \mathbb{R}) = \{A \in \mathbb{R}^{d \times d} : A \text{ regular}\}.$$

Based on the sample $Z_1, \dots, Z_n$, under the group of transformations

$$(Z_1, \dots, Z_n) \mapsto (AZ_1, \dots, AZ_n), \quad A \in \text{GL}(d, \mathbb{R}),$$

the testing problem (2.3) is invariant and Hotelling's T^2 test (2.4) is the uniformly most powerful invariant test of size α ; see Muirhead [19], Theorem 6.3.4.

Under the independent sampling scheme, we choose Hotelling's two-sample T^2 test. The test statistic (modified so as to have an F -distribution) is

$$T'_n = \frac{n(2n-1-d)}{4d(n-1)} (\bar{X}'_n - \bar{Y}'_n)^\top (S_n^{X',Y'})^{-1} (\bar{X}'_n - \bar{Y}'_n), \quad (2.5)$$

where $S^{X'}$ and $S^{Y'}$ are the sample covariance matrices of the sample variables $X'_1, \dots, X'_n$ and $Y'_1, \dots, Y'_n$, and $S_n^{X', Y'} = (S_n^{X'} + S_n^{Y'})/2$. The statistic T'_n has the same distribution as

$$\frac{W'_1/d}{W'_2/(2n-1-d)} \sim F_{d, 2n-1-d}(\delta'), \quad (2.6)$$

where $\delta' = n\mu^\top \Sigma^{-1} \mu/2$ and $W'_1 \sim \chi_d^2(\delta')$, $W'_2 \sim \chi_{2n-1-d}^2$ are independent; see Muirhead [19], Theorems 1.3.6 and 6.3.1. The rejection region of Hotelling's two-sample T^2 test is

$$\{T'_n \geq F_{d, 2n-1-d}^{1-\alpha}\}. \quad (2.7)$$

It is a consistent test of exact size α . In fact, it is the likelihood ratio test. Moreover, under the group of transformations

$$\left(\begin{pmatrix} X'_1 \\ Y'_1 \end{pmatrix}, \dots, \begin{pmatrix} X'_n \\ Y'_n \end{pmatrix} \right) \mapsto \left(\begin{pmatrix} AX'_1 + c \\ AY'_1 + c \end{pmatrix}, \dots, \begin{pmatrix} AX'_n + c \\ AY'_n + c \end{pmatrix} \right), \quad A \in \text{GL}(d, \mathbb{R}), \quad c \in \mathbb{R}^d,$$

the testing problem (2.3) is invariant and the test (2.7) is the uniformly most powerful invariant test of size α ; see Muirhead [19], Theorem 6.3.5.

3. Efficiencies

Let $(\Omega, \mathfrak{A}, \mathfrak{P})$ be a statistical experiment, i.e., a measurable space $(\Omega, \mathfrak{A})$ endowed with a family of probability measures $\mathfrak{P} = \{P_\vartheta : \vartheta \in \Theta\}$ indexed by some parameter ϑ in a parameter set Θ . Fix a subset $\emptyset \neq \Theta_0 \subsetneq \Theta$, set $\Theta_1 = \Theta \setminus \Theta_0$ and assume that the subsets

$$\mathfrak{P}_0 = \{P_\vartheta : \vartheta \in \Theta_0\} \quad \text{and} \quad \mathfrak{P}_1 = \{P_\vartheta : \vartheta \in \Theta_1\}$$

of $\mathfrak{P}$ are disjoint. Consider the question of determining whether the true unknown probability measure P_ϑ is an element of $\mathfrak{P}_0$ (hypothesis H), or an element of $\mathfrak{P}_1$ (alternative K) through a test of the hypotheses

$$\mathcal{H} : \vartheta \in \Theta_0, \quad \mathcal{K} : \vartheta \in \Theta_1. \quad (3.1)$$

Consider treating the testing problem (3.1) by non-randomized statistical tests with rejection regions of the form $\{T \geq c\}$, where $T : (\Omega, \mathfrak{A}) \rightarrow (\mathbb{R}, \mathfrak{B})$ is a test statistic (a map from Ω into $\mathbb{R}$, measurable with respect to $\mathfrak{A}$ and the Borel- σ -algebra $\mathfrak{B}$ on $\mathbb{R}$), and $c \in \mathbb{R}$ is a critical value. Denote such a test briefly by T . In what follows, we deal with sequences of tests $(T_n)_{n \in \mathbb{N}}$ and $(T'_n)_{n \in \mathbb{N}}$ for (3.1), defined in a specific way. We make the following assumptions.

- (a) $(R, \mathfrak{S})$ and $(R', \mathfrak{S}')$ are measurable spaces.
- (b) $\zeta_i : (\Omega, \mathfrak{A}) \rightarrow (R, \mathfrak{S})$, $i \in \mathbb{N}$, and $\zeta'_i : (\Omega, \mathfrak{A}) \rightarrow (R', \mathfrak{S}')$, $i \in \mathbb{N}$, are sequences of random variables. For each $\vartheta \in \Theta$, the ζ_i , $i \in \mathbb{N}$, and the ζ'_i , $i \in \mathbb{N}$, respectively, are independent and identically distributed under $P_\vartheta \in \mathfrak{P}$.

(c) For each $n \in \mathbb{N}$ there exist statistics

$$t_n : (R^n, \mathfrak{S}^n) \rightarrow (\mathbb{R}, \mathfrak{B}) \quad \text{and} \quad t'_n : (R'^n, \mathfrak{S}'^n) \rightarrow (\mathbb{R}, \mathfrak{B})$$

defined on the n -fold product space $(R^n, \mathfrak{S}^n)$ of $(R, \mathfrak{S})$ and the n -fold product space $(R'^n, \mathfrak{S}'^n)$ of $(R', \mathfrak{S}')$, respectively, such that

$$T_n = t_n(\zeta^{(n)}) \quad \text{and} \quad T'_n = t'_n(\zeta'^{(n)}), \quad n \in \mathbb{N},$$

where $\zeta^{(n)} = (\zeta_1, \dots, \zeta_n)$ and $\zeta'^{(n)} = (\zeta'_1, \dots, \zeta'_n)$.

Comparisons of sequences of tests $T = (T_n)_{n \in \mathbb{N}}$ and $T' = (T'_n)_{n \in \mathbb{N}}$ are usually done in the special case $(R, \mathfrak{S}) = (R', \mathfrak{S}')$ and $\zeta_i = \zeta'_i$, $i \in \mathbb{N}$. Then $T_n = t_n(\zeta^{(n)})$ and $T'_n = t'_n(\zeta^{(n)})$ for all $n \in \mathbb{N}$, which means that the sequences of tests T and T' are based on the same sequence of statistical experiments, viz.

$$(R^n, \mathfrak{S}^n, \{P_{\vartheta}^{\zeta^{(n)}} : \vartheta \in \Theta\}), \quad n \geq 1.$$

In our more general setting, the sequences of tests T and T' are based on the different sequences of statistical experiments

$$(R^n, \mathfrak{S}^n, \{P_{\vartheta}^{\zeta^{(n)}} : \vartheta \in \Theta\}), \quad n \geq 1, \quad \text{and} \quad (R'^n, \mathfrak{S}'^n, \{P_{\vartheta}^{\zeta'^{(n)}} : \vartheta \in \Theta\}), \quad n \geq 1,$$

respectively. Of course, it is tacitly assumed that the sets

$$\mathfrak{P}_0^{\zeta^{(n)}} = \{P_{\vartheta}^{\zeta^{(n)}} : \vartheta \in \Theta_0\} \quad \text{and} \quad \mathfrak{P}_1^{\zeta^{(n)}} = \{P_{\vartheta}^{\zeta^{(n)}} : \vartheta \in \Theta_1\}$$

are disjoint, as are the sets

$$\mathfrak{P}_0^{\zeta'^{(n)}} = \{P_{\vartheta}^{\zeta'^{(n)}} : \vartheta \in \Theta_0\} \quad \text{and} \quad \mathfrak{P}_1^{\zeta'^{(n)}} = \{P_{\vartheta}^{\zeta'^{(n)}} : \vartheta \in \Theta_1\}.$$

First, we give a brief overview of the well-known asymptotic efficiency concepts of Hodges–Lehmann, Bahadur, and Pitman, and show their suitability for comparison purposes in our context. For detailed discussions in the traditional setting, see Groeneboom and Oosterhoff [8], Nikitin [20], Puri and Sen [23] and Serfling [26].

3.1. Hodges–Lehmann efficiency

Fix $\alpha \in (0, 1)$. Let $T = (T_n)$ be a sequence of tests (test statistics) T_n , and let (c_n) be the sequence of associated critical values $c_n \in \mathbb{R}$ such that

$$\sup\{P_{\vartheta_0}(T_n \geq c_n) : \vartheta_0 \in \Theta_0\} \leq \alpha \quad \text{for all } n \in \mathbb{N}.$$

For $\vartheta \in \Theta_1$, denote by $\beta_{T,n}(\vartheta) = P_{\vartheta}(T_n < c_n)$ the error probability of the second kind at ϑ . If there exists $d_T(\vartheta) \in (0, \infty)$ such that

$$-d_T(\vartheta) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \beta_{T,n}(\vartheta),$$

$d_T(\vartheta)$ is called the Hodges–Lehmann index of T . Note that this definition differs from that given in the literature, where $2d_T(\vartheta)$ is usually called the Hodges–Lehmann index of T . In what follows, we consider test statistics of the form $T_n = t_n(\zeta^{(n)})$ given above. Then as $n \rightarrow \infty$, $d_T(\vartheta)$ describes the exponential rate of convergence of the error probability of the second kind to zero. If for two given sequences T and T' of competing tests the Hodges–Lehmann indexes $d_T(\vartheta)$ and $d_{T'}(\vartheta)$ exist, the ratio

$$e_{T',T}^{\text{HL}}(\vartheta) = d_{T'}(\vartheta)/d_T(\vartheta)$$

is called the Hodges–Lehmann efficiency of T' with respect to T . In typical cases, the Hodges–Lehmann efficiency does not depend on α . Set

$$K_{\zeta_1}(\vartheta_1, \vartheta_2) = \int \ln \frac{dP_{\vartheta_1}^{\zeta_1}}{dP_{\vartheta_2}^{\zeta_1}} dP_{\vartheta_1}^{\zeta_1}, \quad \vartheta_1, \vartheta_2 \in \Theta,$$

if $P_{\vartheta_1}^{\zeta_1} \ll P_{\vartheta_2}^{\zeta_1}$, and $K_{\zeta_1}(\vartheta_1, \vartheta_2) = \infty$ otherwise. $K_{\zeta_1}(\vartheta_1, \vartheta_2)$ is the Kullback–Leibler information of the distribution $P_{\vartheta_2}^{\zeta_1}$ with respect to the distribution $P_{\vartheta_1}^{\zeta_1}$. Defining

$$K_{\zeta_1}(\Theta_0, \vartheta) = \inf\{K_{\zeta_1}(\vartheta_1, \vartheta), \vartheta_1 \in \Theta_0\},$$

one then has

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \beta_{T,n}(\vartheta) \geq -K_{\zeta_1}(\Theta_0, \vartheta);$$

see, e.g., Kourouklis [13]. Thus, if

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \beta_{T,n}(\vartheta) \leq -K_{\zeta_1}(\Theta_0, \vartheta) < 0,$$

it follows that $d_T(\vartheta) = K_{\zeta_1}(\Theta_0, \vartheta)$. The sequence of tests T is then optimal in the sense of Hodges–Lehmann.

3.2. Bahadur efficiency

For all $t \in \mathbb{R}$, set $G_n(t) = \sup\{P_{\vartheta_0}(T_n \geq t) : \vartheta_0 \in \Theta_0\}$ and define $L_n = G_n(T_n)$. Let $\vartheta \in \Theta_1$. If there exists $c_T(\vartheta) \in (0, \infty)$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln L_n = -c_T(\vartheta) \text{ in } P_{\vartheta}\text{-probability,}$$

$c_T(\vartheta)$ is called the Bahadur slope of T . Note that this definition differs from that given in the literature, where $2c_T(\vartheta)$ is usually called the Bahadur slope of T . For given level α and given $\beta \in (\alpha, 1)$, denote by $N_T = N_T(\alpha, \beta, \vartheta)$ the sample size needed to achieve power β at $\vartheta \in \Theta_1$. Then, in typical cases, $N_T \sim -\ln \alpha / c(\vartheta)$ as $\alpha \rightarrow 0$; see, e.g., Nikitin [20], p. 5. If the Bahadur slopes exist for the sequences of tests T and T' , the ratio

$$e_{T',T}^B(\vartheta) = c_{T'}(\vartheta)/c_T(\vartheta)$$

is called the Bahadur efficiency of T' with respect to T . In view of the reasoning given above, it provides an approximation to the ratio $N_T/N_{T'}$ of sample sizes N_T and $N_{T'}$ required with T and T' . Setting

$$K_{\zeta_1}(\vartheta, \Theta_0) = \inf\{K_{\zeta_1}(\vartheta, \vartheta_2), \vartheta_2 \in \Theta_0\}$$

one then has

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln L_n \geq -K_{\zeta_1}(\vartheta, \Theta_0) \quad P_{\vartheta}\text{-a.s.};$$

see, e.g., Nikitin [20], Theorem 1.2.3. If $c_T(\vartheta) = K_{\zeta_1}(\vartheta, \Theta_0)$, i.e., if for each $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P_{\vartheta} \left\{ \frac{1}{n} \ln L_n + K_{\zeta_1}(\vartheta, \Theta_0) > \varepsilon \right\} = 0,$$

the sequence of tests T is called optimal in the sense of Bahadur.

Let Θ be a metric space. Let $\vartheta_0 \in \partial\Theta_0$, and $(\vartheta_i)_{i \in \mathbb{N}}$ be a sequence of parameters $\vartheta_i \in \Theta_1$, $i \in \mathbb{N}$, such that $\lim_{i \rightarrow \infty} \vartheta_i = \vartheta_0$. If the Bahadur efficiency $e_{T',T}^B(\vartheta_i)$ exists for each $i \in \mathbb{N}$, and if also the limit

$$e_{T',T}^{\text{LB}}\{\vartheta_0, (\vartheta_i)_{i \in \mathbb{N}}\} = \lim_{i \rightarrow \infty} e_{T',T}^B(\vartheta_i)$$

exists, the latter is called the local Bahadur efficiency of T' with respect to T at ϑ_0 with approximating sequence $(\vartheta_i)_{i \in \mathbb{N}}$ in Θ_1 .

3.3. Pitman efficiency

Fix $\alpha \in (0, 1)$. Suppose that Θ is a metric space and that $\partial\Theta_0 \subset \Theta_0$. Take $\vartheta_0 \in \partial\Theta_0$ and a sequence $(\vartheta_n)_{n \in \mathbb{N}}$ of parameters $\vartheta_n \in \Theta_1$ such that $\lim_{n \rightarrow \infty} \vartheta_n = \vartheta_0$. Let us think of $T = (T_n)$ and $T' = (T'_n)_{n \in \mathbb{N}}$ as sequences of statistics for testing, for each given $n \in \mathbb{N}$, the simple hypothesis $\mathcal{H} : \vartheta = \vartheta_0$ against the simple alternative $\vartheta = \vartheta_n$. Let $c_n(\vartheta_0) \in \mathbb{R}$ and $c'_n(\vartheta_0) \in \mathbb{R}$ be two sequences of critical values such that, for every $n \in \mathbb{N}$,

$$P_{\vartheta_0}\{T_n \geq c_n(\vartheta_0)\} \leq \alpha \quad \text{and} \quad P_{\vartheta_0}\{T'_n \geq c'_n(\vartheta_0)\} \leq \alpha$$

and

$$\lim_{n \rightarrow \infty} P_{\vartheta_0}\{T_n \geq c_n(\vartheta_0)\} = \alpha \quad \text{and} \quad \lim_{n \rightarrow \infty} P_{\vartheta_0}\{T'_n \geq c'_n(\vartheta_0)\} = \alpha.$$

If there is a map $h : \mathbb{N} \rightarrow \mathbb{N}$ and $\beta \in (0, 1 - \alpha)$ such that the limits

$$\lim_{n \rightarrow \infty} P_{\vartheta_n} \{T_n < c_n(\vartheta_0)\} = \lim_{n \rightarrow \infty} P_{\vartheta_n} \{T'_{h(n)} < c'_{h(n)}(\vartheta_0)\} = \beta$$

and

$$e_{T', T}^P \{\vartheta_0, (\vartheta_n)_{n \in \mathbb{N}}\} = \lim_{n \rightarrow \infty} n/h(n) \in (0, \infty)$$

exist, the latter is called the Pitman efficiency of T' with respect to T at ϑ_0 with approximating sequence $(\vartheta_n)_{n \in \mathbb{N}}$ in Θ_1 . Often, T_n and T'_n are distribution-free when the hypothesis is true, so that the sequences of critical values do not depend on the specific choice of parameter $\vartheta_0 \in \partial\Theta_0$. Moreover, in typical cases the Pitman efficiency depends neither on α nor on β .

3.4. Volume efficiency

Inspired by the well-known correspondence principle relating (families of) statistical tests and confidence regions, we are led to a different comparison approach based on expected volumes of confidence regions. For a discussion on the expected volume as a criterion for the accuracy of a confidence region, we refer to Pratt [22]. Let us first provide some theoretical background. Fix $\alpha \in (0, 1)$. For some given $p \in \mathbb{N}$, let δ be a map $\delta : \Theta \rightarrow \mathbb{R}^p$. Take $\eta_0 \in \delta(\Theta)$ such that $\Theta_0 = \{\vartheta \in \Theta : \delta(\vartheta) = \eta_0\}$. Suppose that for each $n \in \mathbb{N}$, b_n is a map $b_n : R^n \rightarrow \mathfrak{B}^p$ such that

$$\forall \vartheta \in \Theta \quad \{(z_1, \dots, z_n) \in R^n : \delta(\vartheta) \in b_n(z_1, \dots, z_n)\} \in \mathfrak{S}^n$$

and

$$\forall \vartheta \in \Theta \quad P_{\vartheta}^{(\zeta_1, \dots, \zeta_n)}(\{(z_1, \dots, z_n) \in R^n : \delta(\vartheta) \in b_n(z_1, \dots, z_n)\}) \geq 1 - \alpha.$$

Then, $(b_n)_{n \in \mathbb{N}}$ is a sequence of confidence regions at level $1 - \alpha$ for $\delta(\vartheta)$. Defining, for $n \in \mathbb{N}$ and $(z_1, \dots, z_n) \in R^n$,

$$t_{\eta_0, n}(z_1, \dots, z_n) = 1 \text{ if } \eta_0 \notin b_n(z_1, \dots, z_n) \text{ and } 0 \text{ otherwise,}$$

and setting $T_{\eta_0, n} = t_{\eta_0, n}(\zeta_1, \dots, \zeta_n)$, we have a sequence of non-randomized level α tests with rejection regions

$$\{T_{\eta_0, n} \geq 1\} = \{\eta_0 \notin b_n(\zeta_1, \dots, \zeta_n)\}$$

for the testing problem (3.1). In fact, replacing η_0 by any other $\eta \in \delta(\Theta)$, there is a sequence of non-randomized level α tests (indexed by $\eta \in \delta(\Theta)$) with rejection regions

$$\{T_{\eta, n} \geq 1\} = \{\eta \notin b_n(\zeta_1, \dots, \zeta_n)\}$$

for the testing problems

$$\mathcal{H}_\eta : \delta(\vartheta) = \eta, \quad \mathcal{K}_\eta : \delta(\vartheta) \neq \eta.$$

Denote this family of sequences of tests simply by $T = (T_n)_{n \in \mathbb{N}}$ and set

$$B_{T,n} = b_n^\top(\zeta_1, \dots, \zeta_n) = b_n(\zeta_1, \dots, \zeta_n).$$

Suppose that, for all $n \in \mathbb{N}$,

$$\{(s, z) \in \mathbb{R}^p \times R^n; s \in b_n(z)\} \in \mathfrak{B}^p \otimes \mathfrak{S}^n.$$

This assumption ensures that, for each $n \in \mathbb{N}$, the map $z \mapsto \int_{b_n^\top(z)} ds$, $z \in R^n$, where the symbol ds means integration with respect to the Lebesgue–Borel measure on $\mathfrak{B}^p$, is measurable and that

$$\text{Vol}_{T,n} = \int_{B_{T,n}} ds$$

is a random variable. Let $0 < E_\vartheta(\text{Vol}_{T,n}) < \infty$ for each $\vartheta \in \Theta$, and view $E_\vartheta(\text{Vol}_{T,n})$, $\vartheta \in \Theta$, as a quantification of the localization of $\delta(\vartheta)$ by $B_{T,n}$. Then, for two given such sequences of confidence regions $(b_n^\top)_{n \in \mathbb{N}}$ and $(b_n^{T'})_{n \in \mathbb{N}}$ (two such families of sequences of tests $T = (T_n)_{n \in \mathbb{N}}$ and $T' = (T'_n)_{n \in \mathbb{N}}$), define

$$e_{T',T,n}^{\text{vol}}(\vartheta) = \left\{ \frac{E_\vartheta(\text{Vol}_{T,n})}{E_\vartheta(\text{Vol}_{T',n})} \right\}^2, \quad n \in \mathbb{N}, \quad (3.2)$$

and call this ratio the volume efficiency of the family of tests T' with respect to the family of tests T . The exponent 2 appearing on the right-hand side of (3.2) is applied here to state the conclusions drawn with the various comparison approaches in a simple unified way; the results given below will clarify this. If the limit

$$e_{T',T}^{\text{Lvol}}(\vartheta) = \lim_{n \rightarrow \infty} e_{T',T,n}^{\text{vol}}(\vartheta)$$

exists, it is called the asymptotic volume efficiency of the family of tests T' with respect to the family of tests T .

4. Comparison

First, we compute the efficiencies and show that the Hodges–Lehmann efficiency, the local Bahadur efficiency and the Pitman efficiency are given by

$$E_{T',T}^{\text{HL}}(\eta, \Pi, \Sigma) = E_{T',T}^{\text{LB}}(\eta, \Pi, \Sigma) = E_{T',T}^{\text{P}}(\eta, \Pi, \Sigma) = \eta^\top (2\Sigma)^{-1} \eta / \eta^\top \Lambda^{-1} \eta,$$

respectively, where the parameter $\eta \in \mathbb{R}^d$ is determined by the alternative used for the computation of the efficiency, and that the asymptotic volume efficiency is given by

$$E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) = \det(2\Sigma)^{-1} / \det \Lambda^{-1}.$$

Finally, we use $\lambda_1(\Pi, \Sigma)$ and $\lambda_d(\Pi, \Sigma)$, the extreme eigenvalues of the matrix

$$\Sigma^{-1/2}(\Pi + \Pi^\top)\Sigma^{-1/2}/2,$$

for the characterization of the superiority of one of the sampling designs over the other.

Dealing with the model (2.1) the basic sequences of random variables chosen are

$$\zeta_i = Z_i = X_i - Y_i \quad \text{and} \quad \zeta'_i = (X'_i, Y'_i), \quad i \in \mathbb{N}.$$

The parameter sets are

$$\Theta = \left\{ (\mu, b, \Pi, \Sigma) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d \times d} \times \mathbb{R}^{d \times d} : \begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix} \in \mathcal{P}_{2d} \right\} \quad (4.1)$$

and

$$\Theta_0 = \{ (\mu, b, \Pi, \Sigma) \in \Theta : \mu = 0 \}.$$

For $(\Pi, \Sigma) \in \mathbb{R}^{d \times d} \times \mathbb{R}^{d \times d}$ such that $\begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix} \in \mathcal{P}_{2d}$, define

$$\Lambda = \Lambda(\Pi, \Sigma) = 2\Sigma - (\Pi + \Pi^\top) \quad (4.2)$$

as before. To be specific, for $\vartheta = (\mu, b, \Pi, \Sigma) \in \Theta$ one has

$$P_\vartheta^{\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}} = \mathcal{N}_{2d} \left[\begin{pmatrix} \mu+b \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix} \right],$$

so

$$P_\vartheta^{Z_1} = \mathcal{N}_d(\mu, \Lambda) \quad \text{and} \quad P_\vartheta^{\begin{pmatrix} X'_1 \\ Y'_1 \end{pmatrix}} = \mathcal{N}_{2d} \left[\begin{pmatrix} \mu+b \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} \right].$$

4.1. Computation of the Hodges–Lehmann efficiency

First, for convenience, we state two identities needed in the proof of Proposition 4.1. For $p \in \mathbb{N}$ and positive real η , let $V \sim \chi_p^2(\eta)$. The moment generating function of V is given by

$$\mathbb{E}\{\exp(tV)\} = \exp\left(\frac{t\eta}{1-2t}\right)(1-2t)^{-p/2}, \quad t \in (-\infty, 1/2). \quad (4.3)$$

Let $V_1 \sim \mathcal{N}_d(\nu_1, \Delta_1)$ and $V_2 \sim \mathcal{N}_d(\nu_2, \Delta_2)$, $(\nu_1, \Delta_1), (\nu_2, \Delta_2) \in \mathbb{R}^d \times \mathcal{P}_d$. The Kullback–Leibler information of P^{V_2} with respect to P^{V_1} is

$$\int \ln \frac{dP^{V_1}}{dP^{V_2}} dP^{V_1} = -\frac{1}{2} \ln \det(\Delta_2^{-1} \Delta_1) + \frac{1}{2} (\nu_1 - \nu_2)^\top \Delta_2^{-1} (\nu_1 - \nu_2) + \frac{1}{2} \text{tr}(\Delta_2^{-1} \Delta_1) - \frac{d}{2}. \quad (4.4)$$

Proposition 4.1. *For given fixed alternative specified by the parameter $\vartheta = (\mu, b, \Pi, \Sigma) \in \Theta_1$, the Hodges–Lehmann efficiency of Hotelling’s two-sample T^2 tests T' in independent samples with respect to Hotelling’s T^2 tests T in paired samples is given by*

$$E_{T',T}^{\text{HL}}(\mu, \Pi, \Sigma) = e_{T',T}^{\text{HL}}(\vartheta) = \mu^\top (2\Sigma)^{-1} \mu / \mu^\top \Lambda^{-1} \mu.$$

Proof. For $\vartheta_0 = (0, b_0, \Pi_0, \Sigma_0) \in \Theta_0$, set $\Lambda_0 = 2\Sigma_0 - (\Pi_0 + \Pi_0^\top)$. We then have

$$\begin{aligned} K_{\zeta_1}(\vartheta_0, \vartheta) &= -\frac{1}{2} \ln \det(\Lambda^{-1} \Lambda_0) + \frac{1}{2} \mu^\top \Lambda^{-1} \mu + \frac{1}{2} \text{tr}(\Lambda^{-1} \Lambda_0) - \frac{d}{2} \\ &= \frac{1}{2} \sum_{i=1}^d \{-\ln \lambda_i + (\lambda_i - 1)\} + \frac{1}{2} \mu^\top \Lambda^{-1} \mu, \end{aligned}$$

where $\lambda_1, \dots, \lambda_d$ are the positive real eigenvalues of $\Lambda^{-1} \Lambda_0$. Given that $-\ln x + (x - 1) > 0$ for positive real $x \neq 1$, it follows that

$$K_{\zeta_1}(\Theta_0, \vartheta) = \mu^\top \Lambda^{-1} \mu / 2;$$

see also Kourouklis [13], Example 3.2. Arguing as for the sequence $(\zeta'_n)_{n \in \mathbb{N}}$ below or just refer to Kourouklis [13] to obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \beta_{T,n}(\vartheta) \leq -\mu^\top \Lambda^{-1} \mu / 2.$$

Thus, Hotelling’s T^2 test based on the sample variables Z_i is optimal in the sense of Hodges–Lehmann. The Hodges–Lehmann index is

$$d_T(\vartheta) = \mu^\top \Lambda^{-1} \mu / 2.$$

To get the corresponding result for Hotelling’s two-sample T^2 test, we can call on results of Baringhaus [5], Kallenberg and Kourouklis [10] and Rublík [25], dealing with the Hodges–Lehmann optimality of likelihood ratio tests. For illustration, we prefer to give the derivation for the special case under consideration. Let $\lambda'_1, \dots, \lambda'_d$ be the positive eigenvalues of $\Sigma^{-1} \Sigma_0$. Note that

$$\begin{aligned} K_{\zeta'_1}(\vartheta_0, \vartheta) &= -\ln \det(\Sigma^{-1} \Sigma_0) + \frac{1}{2} (\mu + b - b_0)^\top \Sigma^{-1} (\mu + b - b_0) \\ &\quad + \frac{1}{2} (b - b_0)^\top \Sigma^{-1} (b - b_0) + \text{tr}(\Sigma^{-1} \Sigma_0) - d \\ &= \sum_{i=1}^d \{-\ln \lambda'_i + (\lambda'_i - 1)\} + \frac{1}{2} (\mu + b - b_0)^\top \Sigma^{-1} (\mu + b - b_0) \\ &\quad + \frac{1}{2} (b - b_0)^\top \Sigma^{-1} (b - b_0) \\ &\geq \frac{1}{2} (\mu + b - b_0)^\top \Sigma^{-1} (\mu + b - b_0) + \frac{1}{2} (b - b_0)^\top \Sigma^{-1} (b - b_0) \end{aligned} \tag{4.5}$$

$$\geq \frac{1}{4} \mu^\top \Sigma^{-1} \mu, \tag{4.6}$$

where equality holds in (4.5) and (4.6) if and only if $\Sigma_0 = \Sigma$ and $b_0 = b + \mu/2$, respectively. It follows that

$$K_{\zeta_1}(\Theta_0, \vartheta) = \mu^\top \Sigma^{-1} \mu / 4.$$

Let $c'_n = F_{d, 2n-1-d}^{1-\alpha}$. By using Markov's inequality we deduce from (2.6) and (4.3) that, for each $\tau < 0$ and sufficiently large n ,

$$\begin{aligned} \beta_{T',n}(\vartheta) &= P_\vartheta \left(\tau W'_1 - \frac{\tau d}{2n-1-d} c'_n W'_2 \geq 0 \right) \\ &\leq E_\vartheta \left\{ \exp(\tau W'_1) \right\} E_\vartheta \left\{ \exp \left(- \frac{\tau d}{2n-1-d} c'_n W'_2 \right) \right\} \\ &= \exp \left(\frac{\tau \delta'}{1-2\tau} \right) (1-2\tau)^{-\frac{d}{2}} \left(1 + 2 \frac{\tau d}{2n-1-d} c'_n \right)^{-\frac{2n-1-d}{2}}. \end{aligned}$$

This implies

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \beta_{T',n}(\vartheta) \leq \frac{\tau}{1-2\tau} \frac{1}{2} \mu^\top \Sigma^{-1} \mu,$$

so that as $\tau \rightarrow -\infty$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \beta_{T',n}(\vartheta) \leq -\mu^\top \Sigma^{-1} \mu / 4.$$

Therefore, also Hotelling's two-sample T^2 test based on the sample variables (X'_i, Y'_i) is optimal in the sense of Hodges–Lehmann. The Hodges–Lehmann index is

$$d_{T'}(\vartheta) = \mu^\top \Sigma^{-1} \mu / 4.$$

This proves the assertion. □

4.2. Computation of the Bahadur efficiency

Often it is possible to compute the Bahadur slope in the following way. If there exist a function $b : \Theta_1 \rightarrow \mathbb{R}$ such that

$$\forall \vartheta \in \Theta_1 \quad \lim_{n \rightarrow \infty} T_n = b(\vartheta) \text{ in } P_\vartheta\text{-probability,}$$

an open interval $I \supset \{b(\vartheta) : \vartheta \in \Theta_1\}$, and a continuous function $g : I \rightarrow \mathbb{R}$ such that

$$\forall t \in I \quad \lim_{n \rightarrow \infty} \frac{1}{n} \ln G_n(t) = -g(t),$$

then

$$\forall \vartheta \in \Theta_1 \quad \lim_{n \rightarrow \infty} \frac{1}{n} \ln L_n = -g\{b(\vartheta)\} \text{ in } P_\vartheta\text{-probability.}$$

In this case, the Bahadur slope is given via the function $c_T = g \circ b$; see Bahadur [4], Theorem 7.2. A related result suitable for our purpose is as follows.

Lemma 4.1. *If there exist a function $b : \Theta_1 \rightarrow \mathbb{R}$ such that*

$$\forall \vartheta \in \Theta_1 \quad \lim_{n \rightarrow \infty} T_n = b(\vartheta) \text{ in } P_\vartheta\text{-probability,}$$

an open interval $I \supset \{b(\vartheta) : \vartheta \in \Theta_1\}$, and a function $g : I \rightarrow \mathbb{R}$ continuous from the left such that

$$\forall t \in I \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \ln G_n(t) \leq -g(t),$$

then

$$\forall \vartheta \in \Theta_1 \quad \forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P_\vartheta \left[\frac{1}{n} \ln L_n + g\{b(\vartheta)\} > \varepsilon \right] = 0.$$

In addition, if $g \circ b = K_{\zeta_1}(\cdot, \Theta_0)$, T is optimal in the sense of Bahadur for all $\vartheta \in \Theta_1$.

Proof. Let $\vartheta \in \Theta_1$. For given $\varepsilon > 0$, choose $\delta > 0$ such that

$$b(\vartheta) - \delta \in I \quad \text{and} \quad \tilde{\varepsilon} = \varepsilon - |g\{b(\vartheta) - \delta\} - g\{b(\vartheta)\}| > 0.$$

Since G_n is decreasing, one has $L_n \leq G_n\{b(\vartheta) - \delta\}$ on $\{T_n > b(\vartheta) - \delta\}$. Given that $\limsup_{n \rightarrow \infty} \ln G_n(t)/n \leq -g(t)$ for all $t \in I$ we deduce that

$$\exists k \in \mathbb{N} \quad \forall n \geq k \quad \frac{1}{n} \ln G_n\{b(\vartheta) - \delta\} + g\{b(\vartheta) - \delta\} \leq \tilde{\varepsilon}.$$

This implies

$$\begin{aligned} \exists k \in \mathbb{N} \quad \forall n \geq k \quad \varepsilon &\geq \frac{1}{n} \ln G_n\{b(\vartheta) - \delta\} + g\{b(\vartheta) - \delta\} + |g\{b(\vartheta) - \delta\} - g\{b(\vartheta)\}| \\ &\geq \frac{1}{n} \ln G_n\{b(\vartheta) - \delta\} + g\{b(\vartheta)\}. \end{aligned}$$

Thus, for $n \geq k$,

$$\begin{aligned} P_\vartheta \left[\frac{1}{n} \ln L_n + g\{b(\vartheta)\} > \varepsilon \right] &\leq P_\vartheta \left[\frac{1}{n} \ln G_n\{b(\vartheta) - \delta\} + g\{b(\vartheta)\} > \varepsilon, T_n > b(\vartheta) - \delta \right] \\ &\quad + P_\vartheta \left[\frac{1}{n} \ln L_n + g\{b(\vartheta)\} > \varepsilon, T_n \leq b(\vartheta) - \delta \right] \\ &\leq P_\vartheta \{T_n \leq b(\vartheta) - \delta\}. \end{aligned}$$

The result then follows from the fact that $P_\vartheta\{T_n \leq b(\vartheta) - \delta\} \rightarrow 0$ as $n \rightarrow \infty$. $\square$

Proposition 4.2. *For given fixed alternative specified by the parameter $\vartheta = (\mu, b, \Pi, \Sigma) \in \Theta_1$, the Bahadur efficiency of Hotelling's two-sample T^2 tests T' in independent samples with respect to Hotelling's T^2 tests T in paired samples is given by*

$$e_{T', T}^B(\vartheta) = \ln\{1 + \mu^\top (2\Sigma)^{-1} \mu\} / \ln\{1 + \mu^\top \Lambda^{-1} \mu\}.$$

Proof. Consider the independent samples. We then have

$$\lim_{n \rightarrow \infty} \frac{4d(n-1)}{n(2n-1-d)} T'_n = \mu^\top \Sigma^{-1} \mu \text{ almost surely.}$$

For $\vartheta_0 \in \Theta_0$, Markov's inequality, (2.6) and (4.3) yield

$$\begin{aligned} P_{\vartheta_0} \left\{ \frac{4d(n-1)}{n(2n-1-d)} T'_n \geq t \right\} &= P_{\vartheta_0} \left(\frac{n-1}{2n} W'_1 - \frac{t}{8} W'_2 \geq 0 \right) \\ &\leq E_{\vartheta_0} \left\{ \exp \left(\frac{n-1}{2n} W'_1 \right) \right\} E_{\vartheta_0} \left\{ \exp \left(-\frac{t}{8} W'_2 \right) \right\} \\ &= n^{\frac{d}{2}} (1 + t/4)^{-\frac{2n-1-d}{2}}, \end{aligned}$$

valid for all $t \in [0, \infty)$. This implies

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P_0 \left\{ \frac{4d(n-1)}{n(2n-1-d)} T'_n \geq t \right\} \leq -\ln \left(1 + \frac{t}{4} \right) \leq -\frac{1}{2} \ln \left(1 + \frac{t}{2} \right), \quad t \in [0, \infty).$$

Using (4.4), it turns out that $K_{\zeta'_1}(\vartheta, \Theta_0) = \ln\{1 + \mu^\top (2\Sigma)^{-1} \mu\}/2$ in independent samples. Lemma 4.1 yields the optimality of Hotelling's two-sample T^2 test, i.e.,

$$c_{T'}(\vartheta) = \frac{1}{2} \ln\{1 + \mu^\top (2\Sigma)^{-1} \mu\}.$$

In paired samples, analogous considerations yield that Hotelling's T^2 test based on the sample variables Z_i is optimal in the sense of Bahadur and the Bahadur slope is

$$c_T(\vartheta) = \frac{1}{2} \ln(1 + \mu^\top \Lambda^{-1} \mu);$$

for an alternative approach, see also Koziol [15]. The proposition is thus proven. $\square$

Turning to the local Bahadur efficiency, let us assume that the parameter $\vartheta_0 \in \Theta_0$ and the approximating sequences (ϑ_i) are determined as follows.

(LB) $\vartheta_0 = (0, b, \Pi, \Sigma) \in \Theta_0$, (ϑ_i) is a sequence of elements $\vartheta_i = (\nu_i \eta_i, b_i, \Pi_i, \Sigma_i) \in \Theta_1$, where (ν_i) is a sequence of positive real numbers $\nu_i \rightarrow 0$, (η_i) is a sequence of vectors $\eta_i \in \mathbb{R}^d \setminus \{0\}$ tending to $\eta \in \mathbb{R}^d \setminus \{0\}$, $b_i \rightarrow b$, $\Pi_i \rightarrow \Pi$, $\Sigma_i \rightarrow \Sigma$. Let $\Lambda_i = 2\Sigma_i - (\Pi_i + \Pi_i^\top)$.

Corollary 4.1. *Assume that the condition (LB) is true. Then the local Bahadur efficiency of Hotelling's two-sample T^2 tests T' in independent samples with respect to Hotelling's T^2 tests T in paired samples is given by*

$$E_{T',T}^{\text{LB}}(\eta, \Pi, \Sigma) = e_{T',T}^{\text{LB}}\{\vartheta_0, (\vartheta_i)_{i \in \mathbb{N}}\} = \eta^\top (2\Sigma)^{-1} \eta / \eta^\top \Lambda^{-1} \eta.$$

Proof. By Proposition 4.2, the assertion follows from

$$\lim_{i \rightarrow \infty} \frac{c_{T'}(\vartheta_i)}{c_T(\vartheta_i)} = \lim_{i \rightarrow \infty} \frac{\nu_i^2 \eta_i^\top (2\Sigma_i^{-1}) \eta_i}{\nu_i^2 \eta_i^\top \Lambda_i^{-1} \eta_i} = \frac{\eta^\top (2\Sigma)^{-1} \eta}{\eta^\top \Lambda^{-1} \eta}.$$

□

Note that for the proof of Proposition 4.2, we derived and exploited the Bahadur optimality of the Hotelling tests. General results of Bahadur [3] and Arcones [2] on the Bahadur optimality of likelihood ratio tests may be fruitful in comparing the sampling schemes for other models.

4.3. Computation of the Pitman efficiency

Turning now to the Pitman efficiency, let us assume that the parameter $\vartheta_0 \in \Theta_0$ and the approximating sequences (ϑ_n) are determined as follows.

(P) $\vartheta_0 = (0, b, \Pi, \Sigma) \in \Theta_0$, $\eta \in \mathbb{R}^d \setminus \{0\}$. For $n \in \mathbb{N}$, let $\mu_n = \eta/\sqrt{n} + o(1/\sqrt{n})$, $\vartheta_n = (\mu_n, b_n, \Pi_n, \Sigma_n) \in \Theta_1$, where $b_n \rightarrow b$, $\Pi_n \rightarrow \Pi$, $\Sigma_n \rightarrow \Sigma$. Denote by $h(n)$ the integer part of $n\eta^\top \Lambda^{-1} \eta / \eta^\top (2\Sigma)^{-1} \eta$, $n \in \mathbb{N}$.

Proposition 4.3. *Assume that condition (P) is true. The Pitman efficiency of Hotelling's two-sample T^2 tests T' in independent samples with respect to Hotelling's T^2 tests T in paired samples is given by*

$$E_{T',T}^P(\eta, \Pi, \Sigma) = e_{T',T}^P\{\vartheta_0, (\vartheta_n)_{n \in \mathbb{N}}\} = \eta^\top (2\Sigma)^{-1} \eta / \eta^\top \Lambda^{-1} \eta.$$

Proof. Introduce the independent unit normal random variables $U_1, \dots, U_d, V_1, V_2, \dots$. One then has

$$T_n d \stackrel{\mathcal{D}}{=} \frac{\sum_{i=1}^d \left(U_i + \sqrt{n \mu_n^\top \Lambda_n^{-1} \mu_n / d} \right)^2}{\sum_{i=1}^{n-d} V_i^2 / (n-d)}$$

and

$$T'_{h(n)} d \stackrel{\mathcal{D}}{=} \frac{\sum_{i=1}^d \left(U_i + \sqrt{h(n) \mu_n^\top (2\Sigma_n)^{-1} \mu_n / d} \right)^2}{\sum_{i=1}^{2h(n)-1-d} V_i^2 / (2h(n)-1-d)},$$

where ' $\stackrel{\mathcal{D}}{=}$ ' means equal in distribution. As $n \rightarrow \infty$, almost surely

$$\frac{\sum_{i=1}^d \left(U_i + \sqrt{n \mu_n^\top \Lambda^{-1} \mu_n / d} \right)^2}{\sum_{i=1}^{n-d} V_i^2 / (n-d)} \longrightarrow \sum_{i=1}^d \left(U_i + \sqrt{\eta^\top \Lambda^{-1} \eta / d} \right)^2$$

and

$$\frac{\sum_{i=1}^d \left(U_i + \sqrt{h(n) \mu_n^\top (2\Sigma_n)^{-1} \mu_n / d} \right)^2}{\sum_{i=1}^{2h(n)-1-d} V_i^2 / (2h(n) - 1 - d)} \longrightarrow \sum_{i=1}^d \left(U_i + \sqrt{\eta^\top (2\Sigma)^{-1} \eta / d} \right)^2.$$

Thus

$$\sum_{i=1}^d \left(U_i + \sqrt{\eta^\top \Lambda^{-1} \eta / d} \right)^2 \sim \chi_d^2(\eta^\top \Lambda^{-1} \eta)$$

and

$$\sum_{i=1}^d \left(U_i + \sqrt{\eta^\top (2\Sigma)^{-1} \eta / d} \right)^2 \sim \chi_d^2\{\eta^\top (2\Sigma)^{-1} \eta\}.$$

Let G_b be the distribution function of the $\chi_d^2(b)$ -distribution, $b \in [0, \infty)$. For each $b \in [0, \infty)$, G_b is uniformly continuous and strictly increasing on $[0, \infty)$. With

$$c_n(\vartheta_0) = F_{d,n-d}^{1-\alpha}, \quad c'_n(\vartheta_0) = F_{d,2n-1-d}^{1-\alpha}$$

and $\beta = G_{\eta^\top \Lambda^{-1} \eta}(G_0^{-1}(1-\alpha))$, one has

$$\lim_{n \rightarrow \infty} c_n(\vartheta_0) = \lim_{n \rightarrow \infty} c'_n(\vartheta_0) = G_0^{-1}(1-\alpha)/d$$

as well as

$$\lim_{n \rightarrow \infty} P_{\vartheta_n}\{T_n < c_n(\vartheta_0)\} = \lim_{n \rightarrow \infty} P_{\vartheta_n}\{T'_{h(n)} < c'_{h(n)}(\vartheta_0)\} = \beta$$

and

$$\lim_{n \rightarrow \infty} \frac{n}{h(n)} = \eta^\top (2\Sigma)^{-1} \eta / \eta^\top \Lambda^{-1} \eta.$$

□

For considerations related to that given in the proof of Proposition 4.3, the interested reader is referred to Puri and Sen [23], who deal with the Pitman efficiency concept for (one-dimensional) models with asymptotically non-normal test statistics.

We point out that the results presented in Proposition 4.1, Corollary 4.1 and Proposition 4.3 confirm the findings of Kourouklis [14] and Wieand [29] on the coincidence of the three different efficiency approaches.

4.4. Computation of the volume efficiency

To carry out the comparison based on the volume efficiency, we define $\delta : \Theta \rightarrow \mathbb{R}^d$ by $\delta(\vartheta) = \mu$ for $\vartheta = (\mu, b, \Pi, \Sigma) \in \Theta$. The volume efficiency has to be computed for the two sequences of families of Hotelling tests for the family of testing problems

$$\mathcal{H}_\eta : \delta(\vartheta) = \eta, \quad \mathcal{K}_\eta : \delta(\vartheta) \neq \eta$$

indexed by $\eta \in \mathbb{R}^d$. This means that we have to deal with the associated confidence regions for the subvector μ of ϑ . In paired samples, based on $Z_1, \dots, Z_n$ the convergence region of level $1 - \alpha$ for μ is the confidence ellipsoid given by

$$B_{T,n} = \left\{ \mu \in \mathbb{R}^d : \frac{n(n-d)}{d(n-1)} (\bar{Z}_n - \mu)^\top (S_n^Z)^{-1} (\bar{Z}_n - \mu) \leq F_{d,n-d}^{1-\alpha} \right\}.$$

In independent samples, based on $(X'_1, Y'_1), \dots, (X'_n, Y'_n)$ the convergence region of level $1 - \alpha$ for μ is the confidence ellipsoid given by

$$B_{T',n} = \left\{ \mu \in \mathbb{R}^d : \frac{n(2n-1-d)}{4d(n-1)} (\bar{X}'_n - \bar{Y}'_n - \mu)^\top (S_n^{X',Y'})^{-1} (\bar{X}'_n - \bar{Y}'_n - \mu) \leq F_{d,2n-1-d}^{1-\alpha} \right\}.$$

Proposition 4.4. *The volume efficiency of Hotelling's two-sample T^2 tests T' in independent samples with respect to Hotelling's T^2 tests T in paired samples is given by*

$$e_{T',T,n}^{\text{vol}}(\vartheta) = \frac{(2n-1-d)^d}{2^d(n-d)^d} \frac{(F_{d,n-d}^{1-\alpha})^d}{(F_{d,2n-1-d}^{1-\alpha})^d} \frac{\det \Lambda}{\det \Sigma} \left\{ \frac{\Gamma(\frac{n}{2})\Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2})\Gamma(\frac{2n-1}{2})} \right\}^2.$$

The asymptotic volume efficiency of the sequences is given by

$$E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) = e_{T',T}^{\text{Lvol}}(\vartheta) = \det(2\Sigma)^{-1} / \det \Lambda^{-1}.$$

Proof. Using the well-known formula for volumes of ellipsoids as given, e.g., in Section 9.1 of Königsberger [12], the random volume of the confidence ellipsoid in paired samples is

$$\text{Vol}_{T,n} = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \sqrt{\frac{d^d(n-1)^d}{n^d(n-d)^d} (F_{d,n-d}^{1-\alpha})^d \det S_n^Z}.$$

The random matrix S_n^Z has the Wishart distribution $\mathcal{W}_d(n-1, \Lambda/(n-1))$; see Muirhead [19]. Using Theorem 3.2.15 in Muirhead [19], it is

$$(n-1)^d \frac{\det S_n^Z}{\det \Lambda} \stackrel{\mathcal{D}}{=} \prod_{i=1}^d C_{n-i}$$

with independent random variables $C_{n-d}, \dots, C_{n-1}$, where $C_j \sim \chi_j^2$ for all $j \in \{n-d, \dots, n-1\}$. Then,

$$\begin{aligned} E_\vartheta(\text{Vol}_{T,n}) &= \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \sqrt{\frac{d^d}{n^d(n-d)^d} (F_{d,n-d}^{1-\alpha})^d \det \Lambda \prod_{i=1}^d E_\vartheta(\sqrt{C_{n-i}})} \\ &= \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \sqrt{\frac{d^d}{n^d(n-d)^d} (F_{d,n-d}^{1-\alpha})^d 2^d \det \Lambda \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-d}{2})}}. \end{aligned}$$

In independent samples the random volume of the confidence ellipsoid is given by

$$\text{Vol}_{T',n} = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \sqrt{\frac{4^d d^d (n-1)^d}{n^d (2n-1-d)^d} (F_{d,2n-1-d}^{1-\alpha})^d \det S_n^{X',Y'}}.$$

Since $S_n^{X'}$ and $S_n^{Y'}$ are independent random matrices with the same Wishart distribution $\mathcal{W}_d(n-1, \Sigma/(n-1))$, the random matrix $2S_n^{X',Y'}$ has the Wishart distribution $\mathcal{W}_d(2(n-1), \Sigma/(n-1))$; see, e.g., Muirhead [19], Theorem 3.2.4. One has

$$(n-1)^d \frac{\det(2S_n^{X',Y'})}{\det \Sigma} \stackrel{\mathcal{D}}{=} \prod_{i=1}^d C_{2n-1-i},$$

with independent random variables $C_{2n-1-d}, \dots, C_{2n-2}$, where $C_j \sim \chi_j^2$ for all $j \in \{2n-1-d, \dots, 2n-2\}$. Therefore,

$$\begin{aligned} E_\vartheta(\text{Vol}_{T',n}) &= \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \sqrt{\frac{2^d d^d}{n^d (2n-1-d)^d} (F_{d,2n-1-d}^{1-\alpha})^d \det \Sigma \prod_{i=1}^d E_\vartheta(\sqrt{C_{2n-1-i}})} \\ &= \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \sqrt{\frac{4^d d^d}{n^d (2n-1-d)^d} (F_{d,2n-1-d}^{1-\alpha})^d \det \Sigma \frac{\Gamma(\frac{2n-1}{2})}{\Gamma(\frac{2n-d-1}{2})}} \end{aligned}$$

and

$$\left\{ \frac{E_\vartheta(\text{Vol}_{T,n})}{E_\vartheta(\text{Vol}_{T',n})} \right\}^2 = \frac{(2n-1-d)^d}{2^d (n-d)^d} \frac{(F_{d,n-d}^{1-\alpha})^d}{(F_{d,2n-1-d}^{1-\alpha})^d} \frac{\det \Lambda}{\det \Sigma} \left\{ \frac{\Gamma(\frac{n}{2}) \Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2}) \Gamma(\frac{2n-1}{2})} \right\}^2.$$

Denote by $\chi_{d;1-\alpha}^2$ the quantile of order $1-\alpha$ of the χ_d^2 distribution. Then

$$\lim_{n \rightarrow \infty} F_{d,2n-1-d}^{1-\alpha} = \lim_{n \rightarrow \infty} F_{d,n-d}^{1-\alpha} = \chi_{d;1-\alpha}^2.$$

Furthermore, one has, for each $y \in \mathbb{R}$, $\Gamma(x)/\Gamma(x+y) \sim 1/x^y$ as $x \rightarrow \infty$; see, e.g., Titchmarsh [28], Section 1.87. Consequently,

$$\lim_{n \rightarrow \infty} \left\{ \frac{E_\vartheta(\text{Vol}_{T,n})}{E_\vartheta(\text{Vol}_{T',n})} \right\}^2 = \frac{\det(2\Sigma)^{-1}}{\det \Lambda^{-1}}.$$

□

4.5. Results of the comparison

Let I_d be the $d \times d$ unit matrix. For $\vartheta = (\mu, b, \Pi, \Sigma) \in \Theta$, one has

$$(2\Sigma)^{-1}\Lambda = I_d - \Sigma^{-1}(\Pi + \Pi^\top)/2.$$

Denote by $\Sigma^{-1/2}$ the positive definite square root of Σ^{-1} . The eigenvalues $\lambda_1(\Pi, \Sigma) \leq \dots \leq \lambda_d(\Pi, \Sigma)$ of $\Sigma^{-1}(\Pi + \Pi^\top)/2 = I_d - (2\Sigma)^{-1}\Lambda$ are the same as that of the symmetric matrix $\Sigma^{-1/2}(\Pi + \Pi^\top)\Sigma^{-1/2}/2$. Obviously, $-1 < \lambda_i(\Pi, \Sigma) < +1$ for all $i = 1, \dots, d$. Based on these eigenvalues, there are criteria characterizing the superiority of one of the sampling designs.

Theorem 4.1. *The following statements hold.*

- a) $E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) = \prod_{i=1}^d \{1 - \lambda_i(\Pi, \Sigma)\}.$
- b) $\lambda_1(\Pi, \Sigma) > 0 \Leftrightarrow \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{HL}}(\mu, \Pi, \Sigma) < 1 \Leftrightarrow \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{LB}}(\eta, \Pi, \Sigma) < 1$
 $\Leftrightarrow \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{P}}(\eta, \Pi, \Sigma) < 1.$
 $\lambda_1(\Pi, \Sigma) > 0 \Rightarrow E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) < 1.$
- c) $\lambda_1(\Pi, \Sigma) \geq 0 \Leftrightarrow \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{HL}}(\mu, \Pi, \Sigma) \leq 1 \Leftrightarrow \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{LB}}(\eta, \Pi, \Sigma) \leq 1$
 $\Leftrightarrow \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{P}}(\eta, \Pi, \Sigma) \leq 1.$
 $\lambda_1(\Pi, \Sigma) \geq 0 \Rightarrow E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) \leq 1.$
- d) $\lambda_d(\Pi, \Sigma) \leq 0 \Leftrightarrow \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{HL}}(\mu, \Pi, \Sigma) \geq 1 \Leftrightarrow \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{LB}}(\eta, \Pi, \Sigma) \geq 1$
 $\Leftrightarrow \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{P}}(\eta, \Pi, \Sigma) \geq 1.$
 $\lambda_d(\Pi, \Sigma) \leq 0 \Rightarrow E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) \geq 1.$
- e) $\lambda_d(\Pi, \Sigma) < 0 \Leftrightarrow \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{HL}}(\mu, \Pi, \Sigma) > 1 \Leftrightarrow \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{LB}}(\eta, \Pi, \Sigma) > 1$
 $\Leftrightarrow \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} E_{T',T}^{\text{P}}(\eta, \Pi, \Sigma) > 1.$
 $\lambda_d(\Pi, \Sigma) < 0 \Rightarrow E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) > 1.$

Proof. Note that the eigenvalues of $(2\Sigma)^{-1}\Lambda = I_d - \Sigma^{-1}(\Pi + \Pi^\top)/2$ are $1 - \lambda_1(\Pi, \Sigma), \dots, 1 - \lambda_d(\Pi, \Sigma)$. Thus,

$$E_{T',T}^{\text{Lvol}}(\Pi, \Sigma) = \det(2\Sigma)^{-1}\Lambda = \prod_{i=1}^d \{1 - \lambda_i(\Pi, \Sigma)\}.$$

This proves statement a) of the theorem. Denote by $\Lambda^{1/2} \in \mathcal{P}_d$ the square root of Λ . Define

$$K(\Lambda, \Sigma) = \Lambda^{1/2}(2\Sigma)^{-1}\Lambda^{1/2}.$$

$K(\Lambda, \Sigma)$ has the same eigenvalues as $(2\Sigma)^{-1}\Lambda$. Set $\xi = \Lambda^{-1/2}\nu$ for $\nu \in \mathbb{R}^d \setminus \{0\}$. Then,

$$\frac{\nu^\top (2\Sigma)^{-1}\nu}{\nu^\top \Lambda^{-1}\nu} = \frac{\xi^\top K(\Lambda, \Sigma)\xi}{\xi^\top \xi}$$

and so

$$1 - \lambda_d(\Pi, \Sigma) = \inf_{\xi \neq 0} \frac{\xi^\top K(\Lambda, \Sigma) \xi}{\xi^\top \xi} = \inf_{\nu \neq 0} \frac{\nu^\top (2\Sigma)^{-1} \nu}{\nu^\top \Lambda^{-1} \nu}$$

and

$$1 - \lambda_1(\Pi, \Sigma) = \sup_{\xi \neq 0} \frac{\xi^\top K(\Lambda, \Sigma) \xi}{\xi^\top \xi} = \sup_{\nu \neq 0} \frac{\nu^\top (2\Sigma)^{-1} \nu}{\nu^\top \Lambda^{-1} \nu}.$$

This proves assertions b)–e) of the theorem with the appropriate change of notations. $\square$

5. Pretests for superiority

Assume that observations (from earlier studies, for example) obtained with the paired sampling scheme are available. Given a choice between paired or independent survey sampling in future research, any indication as to which design is preferable would be welcome. In view of the results just obtained, this motivates treating the testing problems

$$\mathcal{H}_1 : \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} \leq 1, \quad \mathcal{K}_1 : \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} > 1 \quad (5.1)$$

and

$$\mathcal{H}_2 : \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} \geq 1, \quad \mathcal{K}_2 : \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} < 1 \quad (5.2)$$

based on observations of the paired sample variables $(X_1, Y_1), \dots, (X_n, Y_n)$, say, where

$$P_{\vartheta}^{(X_1, Y_1)} = \mathcal{N}_{2d} \left[\begin{pmatrix} \mu^+ \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix} \right] \quad (5.3)$$

and $\vartheta = (\mu, b, \Pi, \Sigma)$ is an element of the parameter set Θ defined in (4.1). As before, put $\Xi = \begin{pmatrix} \Sigma & \Pi \\ \Pi^\top & \Sigma \end{pmatrix}$. To describe the test statistic to be suggested, we start with the empirical covariance matrix S_n of $(X_1, Y_1), \dots, (X_n, Y_n)$ given in the partitioned form

$$S_n = \begin{pmatrix} S_n^{11} & S_n^{12} \\ S_n^{21} & S_n^{22} \end{pmatrix}$$

with random $d \times d$ matrices $S_n^{11}, S_n^{12}, S_n^{21}$ and S_n^{22} . For any $k \in \mathbb{N}$ denote by $\text{vec} : \mathbb{R}^{k \times k} \rightarrow \mathbb{R}^{k^2}$ the vector operator defined by

$$\text{vec}(M) = (m_{11}, \dots, m_{k1}, \dots, m_{k1}, \dots, m_{kk})^\top, \quad M = (m_{ij})_{1 \leq i, j \leq k} \in \mathbb{R}^{k \times k}.$$

Set $p = \text{vec}(\Pi)$ and define the estimator $\hat{p}_n = \text{vec}(S_n^{12})$. For a symmetric matrix $M \in \mathbb{R}^{k \times k}$ we use the common notation $M > 0$ ($M \geq 0$, $M \leq 0$, $M < 0$) to express that M is positive definite (positive semidefinite, negative semidefinite, negative definite). One then has

$$\forall_{\mu \in \mathbb{R}^d} \quad \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} \begin{cases} < \\ \leq \\ \geq \\ > \end{cases} 1 \quad \Leftrightarrow \quad (2\Sigma)^{-1} - \Lambda^{-1} \begin{cases} < \\ \leq \\ \geq \\ > \end{cases} 0 \quad \Leftrightarrow \quad 2\Sigma - \Lambda \begin{cases} > \\ \geq \\ \leq \\ < \end{cases} 0.$$

For the last equivalence, see, e.g., Gruber [9], Theorem 8.8. Denoting by $\ell_1(p) \leq \dots \leq \ell_d(p)$ the eigenvalues of the symmetrized cross-covariance matrix

$$H(p) = \frac{1}{2} (\Pi + \Pi^\top) = \Sigma - \frac{1}{2} \Lambda,$$

it is seen that the testing problems (5.1) and (5.2) can be stated as

$$\mathcal{H}_1 : \ell_1(p) \geq 0, \quad \mathcal{K}_1 : \ell_1(p) < 0 \quad (5.4)$$

and

$$\mathcal{H}_2 : \ell_d(p) \leq 0, \quad \mathcal{K}_2 : \ell_d(p) > 0. \quad (5.5)$$

Denote by $\ell_{1,n} = \ell_1(\hat{p}_n) \leq \dots \leq \ell_{d,n} = \ell_d(\hat{p}_n)$ the eigenvalues of symmetrized sample cross-covariance matrix

$$H(\hat{p}_n) = \frac{1}{2} (S_n^{12} + S_n^{21}) = \frac{1}{2} (S_n^{12} + (S_n^{12})^\top).$$

The first tests to be suggested are based on these empirical eigenvalues. In order to determine the asymptotic properties of these tests, results on the behavior of these eigenvalues are needed. For $k \in \mathbb{N}$, let $\mathcal{S}_k := \{A \in \mathbb{R}^{k \times k}; A \text{ is symmetric}\}$. For $i = 1, \dots, k$, denote by $\mathcal{S}_k^i \subset \mathcal{S}_k$ the subset of matrices $M \in \mathcal{S}_k$ with unique i th eigenvalue. Define the operator $\text{vech} : \mathcal{S}_k \rightarrow \mathbb{R}^{k(k+1)/2}$ by

$$\text{vech}(M) = (m_{11}, \dots, m_{k1}, m_{22}, \dots, m_{k2}, \dots, m_{kk})^\top, \quad M = (m_{ij})_{1 \leq i, j \leq k} \in \mathcal{S}_k,$$

and denote by $D \in \mathbb{R}^{k^2 \times k(k+1)/2}$ the duplication matrix, i.e., the unique matrix with the property

$$\forall_{M \in \mathcal{S}_k} \quad \text{vec}(M) = D \text{vech}(M);$$

see Magnus and Neudecker [18], Section 3.8. D has rank $k(k+1)/2$. Putting $G = (D^\top D)^{-1} D^\top$ it is

$$\forall_{M \in \mathcal{S}_k} \quad \text{vech}(M) = G \text{vec}(M).$$

Identifying $\mathcal{S}_k$ with $\mathbb{R}^{k(k+1)/2}$, it is seen that $\mathcal{S}_k^i$ is an open subset of $\mathcal{S}_k$. In fact, by the Wielandt-Hoffman theorem the vector $(l_1(M), \dots, l_k(M))$ of the ordered eigenvalues $l_1(M) \leq \dots \leq l_k(M)$ of M is a continuous function of $M \in \mathcal{S}_k$. See Wilkinson [30], p. 104. In what follows, an additional smoothness property of the eigenvalues as functions of M is needed.

Lemma 5.1. (Magnus [17]) *Let $i \in \{1, \dots, k\}$ be fixed. Let $b_{0,i}$ be a normalized eigenvector associated with the unique eigenvalue $l_{0,i}$ of $M_0 \in \mathcal{S}_k^i$. Then, there are measurable functions $l_i : \mathcal{S}_k \rightarrow \mathbb{R}$, $b_i : \mathcal{S}_k \rightarrow \mathbb{R}^k$ and an open neighborhood $U \subset \mathcal{S}_k^i$ of M_0 , such that l_i and b_i are continuously differentiable on U and*

$$\begin{aligned} l_i(M_0) &= l_{0,i}, \quad b_i(M_0) = b_{0,i}, \\ \forall M \in U \quad Mb_i(M) &= l_i(M)b_i(M), \quad \forall M \in \mathcal{S}_k : b_i(M)^\top b_i(M) = I_k. \end{aligned}$$

The differential of l_i for $M \in U$ is given by

$$\frac{\partial l_i(M)}{\partial \text{vech}(M)} = (b_i(M)^\top \otimes b_i(M)^\top)D, \quad M \in U, \quad (5.6)$$

where $\otimes$ denotes the Kronecker product.

For the proof we refer to Magnus [17]. In what follows, $\rightsquigarrow$ means convergence in distribution.

Lemma 5.2. *Let $M_n, n \in \mathbb{N}$, be symmetric random $k \times k$ matrices and let $M \in \mathcal{S}_k^i$ be a matrix such that*

$$\sqrt{n} \{ \text{vech}(M_n) - \text{vech}(M) \} \rightsquigarrow \mathcal{N}_{k(k+1)/2}(0, \Delta) \text{ as } n \rightarrow \infty$$

with a covariance matrix $\Delta \in \mathbb{R}^{k(k+1)/2 \times k(k+1)/2}$. Then, as $n \rightarrow \infty$,

$$\sqrt{n} \{ l_i(M_n) - l_i(M) \} \rightsquigarrow \mathcal{N} \left[0, \{ b_i(M)^\top \otimes b_i(M)^\top \} D \Delta D^\top \{ b_i(M) \otimes b_i(M) \} \right].$$

Proof. This follows from Lemma 5.1 by using the delta method. $\square$

Let us come back to the model (5.3) posed at the beginning of this section. In what follows, to deal with the testing problem (5.4) and the testing problem (5.5) we assume that the parameter space Θ is restricted in the way that $H(p) \in \mathcal{S}_d^1$ and that $H(p) \in \mathcal{S}_d^d$, respectively. It holds that

$$\sqrt{n} \{ \text{vech}(S_n) - \text{vech}(\Xi) \} \rightsquigarrow \mathcal{N}_{2d(2d+1)/2}[0, \text{Iss}(\Xi)] \text{ as } n \rightarrow \infty, \quad (5.7)$$

where $\text{Iss}(\Xi) \in \mathcal{S}_{2d(2d+1)/2}$ is the Isserlis matrix of Ξ . With $\Xi = (\xi_{ij})_{1 \leq i, j \leq 2d}$ the entries of the (symmetric, positive semidefinite matrix) $\text{Iss}(\Xi)$ are given by

$$\xi_{ik}\xi_{jl} + \xi_{il}\xi_{jk}, \quad 1 \leq i \leq j \leq 2d, \quad 1 \leq k \leq l \leq 2d.$$

See Anderson [1], Theorem 3.4.3 and Theorem 3.4.4. Due to our assumption that X_1 and Y_1 have the same covariance matrix Σ , the Isserlis matrix depends on Π and Σ only. So

in what follows it is also written as $\text{Iss}(\Pi, \Sigma)$. From Roverato and Whittaker [24], we know that

$$\det \text{Iss}(\Xi) = 2^{2d} (\det \Xi)^{2d+1}$$

and hence $\text{Iss}(\Xi)$ is positive definite. Denote by $\mathcal{J}_H(p) \in \mathbb{R}^{d(d+1)/2 \times d^2}$ the Jacobian of the continuously differentiable map $\text{vech}(H(\cdot)) : \mathbb{R}^{d^2} \rightarrow \mathbb{R}^{d(d+1)/2}$ at p . The matrix $\mathcal{J}_H(p)$ has full row rank. Let $M \in \mathcal{S}_{2d}$ be partitioned as

$$M = \begin{pmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{pmatrix} \quad (5.8)$$

with submatrices $M^{11}, M^{12}, M^{21}, M^{22} \in \mathbb{R}^{d \times d}$. Denote by $B \in \mathbb{R}^{d^2 \times 2d(2d+1)/2}$ the full row rank matrix that is uniquely determined by

$$\forall_{M \in \mathcal{S}_{2d}} \quad B \text{vech}(M) = \text{vec}(M_{12}).$$

Then $V(\Pi, \Sigma) = B \text{Iss}(\Pi, \Sigma) B^\top > 0$.

Proposition 5.1. *For given $i \in \{1, d\}$, let $\beta_i(p)$ be a normalized eigenvector associated with the eigenvalue $\ell_i(p)$ of $H(p)$. For each fixed $\vartheta \in \Theta$ such that $H(p) \in \mathcal{S}_d^i$ we have that, as $n \rightarrow \infty$,*

$$\sqrt{n} \{ \ell_i(\hat{p}_n) - \ell_i(p) \} \rightsquigarrow \mathcal{N}[0, N_i(\Pi, \Sigma)], \quad (5.9)$$

where

$$N_i(\Pi, \Sigma) = \{ \beta_i(p)^\top \otimes \beta_i(p)^\top \} D \mathcal{J}_H(p) V(\Pi, \Sigma) \mathcal{J}_H(p)^\top D^\top \{ \beta_i(p) \otimes \beta_i(p) \} > 0.$$

Moreover,

$$\lim_{n \rightarrow \infty} N_i \left\{ S_n^{12}, \frac{1}{2} (S_n^{11} + S_n^{22}) \right\} = N_i(\Pi, \Sigma) \text{ in } P_\vartheta\text{-probability.} \quad (5.10)$$

Proof. From (5.7) it follows that

$$\sqrt{n} \{ B \text{vech}(S_n) - B \text{vech}(\Xi) \} \rightsquigarrow \mathcal{N}_{d^2} [0, B \text{Iss}(\Pi, \Sigma) B^\top],$$

equivalently,

$$\sqrt{n} (\hat{p}_n - p) \rightsquigarrow \mathcal{N}_{d^2} [0, V(\Pi, \Sigma)]. \quad (5.11)$$

The delta method gives

$$\sqrt{n} \{ (\text{vech}(H(\hat{p}_n)) - \text{vech}(H(p))) \} \rightsquigarrow \mathcal{N}_{d(d+1)/2} [0, \mathcal{J}_H(p) V(\Pi, \Sigma) \mathcal{J}_H(p)^\top].$$

Note, that $(\beta_i(p)^\top \otimes \beta_i(p)^\top) D$ does not vanish. Use Lemmas 5.1 and 5.2 to see that the assertions (5.9) and (5.10) hold. $\square$

Set $N_{i,n} = N_i \{S_n^{12}, (S_n^{11} + S_n^{22})/2\}$. Under the conditions of Proposition 5.1,

$$\lim_{n \rightarrow \infty} P_{\vartheta}(N_{i,n} > 0) = 1.$$

We define the test statistics

$$T_{1,n} = -\frac{\sqrt{n} \ell_{1,n}}{\sqrt{N_{1,n}}} \text{ if } N_{1,n} > 0 \text{ and } 0 \text{ otherwise,} \quad (5.12)$$

and

$$T_{2,n} = \frac{\sqrt{n} \ell_{d,n}}{\sqrt{N_{d,n}}} \text{ if } N_{d,n} > 0 \text{ and } 0 \text{ otherwise.} \quad (5.13)$$

As a simple consequence of Proposition 5.1, we have the following result.

Theorem 5.1. a) Assume that the parameter space is restricted in such a way that $H(p) \in \mathcal{S}_d^1$. Then

$$\lim_{n \rightarrow \infty} P_{\vartheta}(T_{1,n} \geq u_{1-\alpha}) = \begin{cases} 1, & \text{if } \ell_1(p) < 0, \\ \alpha, & \text{if } \ell_1(p) = 0, \\ 0, & \text{if } \ell_1(p) > 0. \end{cases}$$

b) Assume that the parameter space is restricted in such a way that $H(p) \in \mathcal{S}_d^d$. Then

$$\lim_{n \rightarrow \infty} P_{\vartheta}(T_{2,n} \geq u_{1-\alpha}) = \begin{cases} 1, & \text{if } \ell_d(p) > 0, \\ \alpha, & \text{if } \ell_d(p) = 0, \\ 0, & \text{if } \ell_d(p) < 0. \end{cases}$$

Theorem 5.1 ensures that for the testing problems (5.1) and (5.2) the corresponding tests with rejection regions $\{T_{1,n} \geq u_{1-\alpha}\}$ and $\{T_{2,n} \geq u_{1-\alpha}\}$ are asymptotically of given size $\alpha \in (0, 1)$. Moreover, the tests are consistent.

Remark 5.1. The test statistics $T_{1,n}$ and $T_{2,n}$ are built from the estimators

$$N_{i,n} = N_i \left\{ S_n^{12}, \frac{1}{2} (S_n^{11} + S_n^{22}) \right\}$$

of the unknown variances $N_i(\Xi) = N_i(\Pi, \Sigma)$ of the limit normal distributions of $\ell_{i,n} = \ell_i(\hat{p}_n)$, $i = 1, d$, given in Proposition 5.1. Obviously, $N_i(\Xi)$ is consistently estimated also by $N_i(S_n)$. Thus, replacing $N_{i,n}$ by $N_i(S_n)$, Theorem 5.1 still holds.

Other related consistent tests can be obtained in the following way. The testing problems (5.1) and (5.2) can be also stated as

$$\mathcal{H}_1 : \lambda_1(\Pi, \Sigma) \geq 0, \quad \mathcal{K}_1 : \lambda_1(\Pi, \Sigma) < 0$$

and

$$\mathcal{H}_2 : \lambda_d(\Pi, \Sigma) \leq 0, \quad \mathcal{K}_2 : \lambda_d(\Pi, \Sigma) > 0.$$

Suitable estimators of these eigenvalues are the eigenvalues $\lambda_{1,n} \leq \dots \leq \lambda_{d,n}$ of the matrix

$$(S_n^{11} + S_n^{22})^{-1} (S_n^{12} + S_n^{21}).$$

Let us speak of the sample canonical correlation type coefficients $\lambda_{i,n}$. For the latter, arguing as above one can derive an asymptotic distributional result corresponding to that for $\ell_{i,n} = \ell_i(\hat{p}_n)$ given in Proposition 5.1. For $\vartheta = (\mu, b, \Pi, \Sigma) \in \Theta$ with

$$\frac{1}{2} \Sigma^{-1/2} (\Pi + \Pi^\top) \Sigma^{-1/2} \in \mathcal{S}_d^i,$$

the variance $\tilde{N}_i(\Xi) = \tilde{N}_i(\Pi, \Sigma)$ of the limit normal distribution depends continuously on (Π, Σ) . Estimate $\tilde{N}_i(\Pi, \Sigma)$ by

$$\tilde{N}_{i,n} = \tilde{N}_i \left\{ S_n^{12}, \frac{1}{2} (S_n^{11} + S_n^{22}) \right\}$$

and define the test statistics

$$\tilde{T}_{1,n} = -\frac{\sqrt{n} \lambda_{1,n}}{\sqrt{\tilde{N}_{1,n}}} \text{ if } \tilde{N}_{1,n} > 0 \text{ and } 0 \text{ otherwise,} \quad (5.14)$$

and

$$\tilde{T}_{2,n} = \frac{\sqrt{n} \lambda_{d,n}}{\sqrt{\tilde{N}_{d,n}}} \text{ if } \tilde{N}_{d,n} > 0 \text{ and } 0 \text{ otherwise.} \quad (5.15)$$

It is seen that if the parameter space is restricted in the way that $\Sigma^{-1/2}(\Pi + \Pi^\top)\Sigma^{-1/2}/2 \in \mathcal{S}_d^1$, then

$$\lim_{n \rightarrow \infty} P_\vartheta(\tilde{T}_{1,n} \geq u_{1-\alpha}) = \begin{cases} 1, & \text{if } \lambda_1(\Pi, \Sigma) < 0, \\ \alpha, & \text{if } \lambda_1(\Pi, \Sigma) = 0, \\ 0, & \text{if } \lambda_1(\Pi, \Sigma) > 0. \end{cases} \quad (5.16)$$

and that if the parameter space is restricted in the way that $\Sigma^{-1/2}(\Pi + \Pi^\top)\Sigma^{-1/2}/2 \in \mathcal{S}_d^d$, then

$$\lim_{n \rightarrow \infty} P_\vartheta(\tilde{T}_{2,n} \geq u_{1-\alpha}) = \begin{cases} 1, & \text{if } \lambda_d(\Pi, \Sigma) > 0, \\ \alpha, & \text{if } \lambda_d(\Pi, \Sigma) = 0, \\ 0, & \text{if } \lambda_d(\Pi, \Sigma) < 0. \end{cases} \quad (5.17)$$

Thus, the tests with rejection regions $\{\tilde{T}_{1,n} \geq u_{1-\alpha}\}$ and $\{\tilde{T}_{2,n} \geq u_{1-\alpha}\}$ are consistent and asymptotically of given size $\alpha \in (0, 1)$. Of course, it is possible to replace $\tilde{N}_{i,n}$ by $\tilde{N}_i(S_n)$, $i = 1, d$; the assertions (5.16) and (5.17) still hold; see Remark 5.1.

In view of the volume efficiency based dominance result we are led to consider also the testing problems

$$\mathcal{H}_3 : \det((2\Sigma)^{-1}\Lambda) \leq 1, \quad \mathcal{K}_3 : \det((2\Sigma)^{-1}\Lambda) > 1 \quad (5.18)$$

and

$$\mathcal{H}_4 : \det((2\Sigma)^{-1}\Lambda) \geq 1, \quad \mathcal{K}_4 : \det((2\Sigma)^{-1}\Lambda) < 1. \quad (5.19)$$

On writing the elements $M \in \mathcal{P}_{2d}$ in the canonical partitioned form shown in (5.8), we define the function $\mathbb{D} : \mathcal{P}_{2d} \rightarrow \mathbb{R}^{d \times d}$ by

$$\mathbb{D}(M) = I_d - (M^{11} + M^{22})^{-1} (M^{12} + M^{21}), \quad M \in \mathcal{P}_{2d}.$$

One has

$$\det \mathbb{D}(S_n) = \det \left(I_d - (S_n^{11} + S_n^{22})^{-1} (S_n^{12} + S_n^{21}) \right) = \prod_{i=1}^d (1 - \lambda_{i,n}),$$

which is a natural estimator of

$$\det((2\Sigma)^{-1}\Lambda) = \det(I_d - (2\Sigma)^{-1}(\Pi + \Pi^\top)) = \prod_{i=1}^d (1 - \lambda_i(\Pi, \Sigma)).$$

We aim to suggest asymptotically studentized versions as test statistics. To this end, we first derive the limit distribution of $\det \mathbb{D}(S_n)$. Note that $\det \mathbb{D}(M) > 0$ for $M \in \mathcal{P}_{2d}$. By symmetry of M , regard $\mathbb{D}$ as a matrix-valued function of $\text{vech}(M)$. In this sense, $\mathbb{D}$ is differentiable, its Jacobian $\mathcal{J}_{\mathbb{D}}(M) \in \mathbb{R}^{d(d+1)/2 \times 2d(2d+1)/2}$, $M \in \mathcal{P}_{2d}$, is continuous and of full row rank at all $M \in \mathcal{P}_{2d}$. At regular $A_0 \in \mathcal{S}_d$, the gradient (row vector) $\mathcal{J}_{\det}(A_0) \in \mathbb{R}^{d(d+1)/2}$ of the function $\det A$, $A \in \mathcal{S}_d$, also regarded as a function of $\text{vech}(A)$, $A \in \mathcal{S}_d$, does not vanish. It follows that

$$v(M) = \mathcal{J}_{\det}(\mathbb{D}(M)) \mathcal{J}_{\mathbb{D}}(M) \text{Iss}(M) \mathcal{J}_{\mathbb{D}}(M)^\top \mathcal{J}_{\det}(\mathbb{D}(M))^\top$$

is positive for each $M \in \mathcal{P}_{2d}$. Due to (5.7), by applying the delta method and using the continuity of v we obtain the following result.

Proposition 5.2. *For each $\vartheta = (\mu, b, \Pi, \Sigma)$ with associated matrix Ξ it holds that, as $n \rightarrow \infty$,*

$$\sqrt{n} \{ \det \mathbb{D}(S_n) - \det \mathbb{D}(\Xi) \} \rightsquigarrow \mathcal{N}[0, v(\Xi)] \quad (5.20)$$

Moreover,

$$\lim_{n \rightarrow \infty} v(S_n) = v(\Xi) \text{ in } P_\vartheta\text{-probability.}$$

Putting $L_n = v(S_n)$, we define the test statistics

$$T_{3,n} = \frac{\sqrt{n} \{\det \mathbb{D}(S_n) - 1\}}{\sqrt{L_n}} \quad \text{and} \quad T_{4,n} = -\frac{\sqrt{n} \{\det \mathbb{D}(S_n) - 1\}}{\sqrt{L_n}}.$$

For the testing problems (5.18) and (5.19), we suggest the test with rejection region $\{T_{3,n} > u_{1-\alpha}\}$ and $\{T_{4,n} > u_{1-\alpha}\}$, respectively.

As consequence of Proposition 5.2 we get the following result.

Theorem 5.2. a)

$$\lim_{n \rightarrow \infty} P_{\vartheta}(T_{3,n} \geq u_{1-\alpha}) = \begin{cases} 0, & \text{if } \det \mathbb{D}(\Xi) < 1, \\ \alpha, & \text{if } \det \mathbb{D}(\Xi) = 1, \\ 1, & \text{if } \det \mathbb{D}(\Xi) > 1. \end{cases}$$

b)

$$\lim_{n \rightarrow \infty} P_{\vartheta}(T_{4,n} \geq u_{1-\alpha}) = \begin{cases} 0, & \text{if } \det \mathbb{D}(\Xi) > 1, \\ \alpha, & \text{if } \det \mathbb{D}(\Xi) = 1, \\ 1, & \text{if } \det \mathbb{D}(\Xi) < 1. \end{cases}$$

Thus the tests are asymptotically of given size $\alpha \in (0, 1)$. Moreover, the tests are consistent.

Remark 5.2. The test statistics $T_{3,n}$ and $T_{4,n}$ are built with the estimator $L_n = v(S_n)$ of the unknown covariance matrix $v(\Xi)$ of the limit normal distributions of $\det \mathbb{D}(S_n)$ given in Proposition 5.20. $v(\Xi) = v(\Pi, \Sigma)$ is a function of Π and Σ only. It is consistently estimated also by $\tilde{L}_n = v(S_n^{12}, (S_n^{11} + S_n^{22})/2)$. Replacing, L_n by $\tilde{L}_n$, Theorem 5.2 still hold.

The special case $d = 1$ deserves special attention. In fact, if $d = 1$, the testing problems (5.1) and (5.18) as well as the testing problems (5.2) and (5.19) coincide. By noting that $\det((2\Sigma)^{-1}\Lambda) = 1 - \rho$, where ρ is the correlation coefficient of X_1 and Y_1 , the testing problems can be simply written as $\mathcal{H}_3 : \rho \geq 0$, $\mathcal{K}_3 : \rho < 0$ and $\mathcal{H}_4 : \rho \leq 0$, $\mathcal{K}_4 : \rho > 0$, where ρ denotes the correlation coefficient of X_1 and Y_1 . In what follows, let us consider the latter testing problem only. Setting

$$W_n = \frac{2S_n^{12}}{S_n^{11} + S_n^{22}},$$

one gets

$$T_{4,n} = \sqrt{n} \frac{W_n}{\sqrt{L_n}}.$$

One can then apply (5.20) to see that, as $n \rightarrow \infty$,

$$\sqrt{n}(W_n - \rho) \rightsquigarrow \mathcal{N}[0, (1 - \rho^2)^2].$$

It is worth noting that W_n has the same limit distribution as the empirical correlation coefficient $R_n = S_n^{12}/\sqrt{S_n^{11}S_n^{22}}$; see Muirhead [19], p.159. In fact, due to the assumption $\text{var}(X_1) = \text{var}(Y_1)$ we have two consistent estimators W_n and R_n for ρ . Estimating ρ by R_n , one has $L_n = v(S_n) = (1 - R_n^2)^2$, so

$$T_{4,n} = \sqrt{n} \frac{W_n}{1 - R_n^2}.$$

Replacing R_n by W_n we obtain the related test statistic

$$\tilde{T}_{4,n} = \sqrt{n} \frac{W_n}{1 - W_n^2}.$$

The test based on $\tilde{T}_{4,n}$ with the rejection region $\{\tilde{T}_{4,n} > u_{1-\alpha}\}$ behaves asymptotically as that based on $T_{4,n}$, i.e.,

$$\lim_{n \rightarrow \infty} P_{\vartheta}(\tilde{T}_{4,n} \geq u_{1-\alpha}) = \begin{cases} 0, & \text{if } \rho < 0, \\ \alpha, & \text{if } \rho = 0, \\ 1, & \text{if } \rho > 0. \end{cases}$$

However, a better even non-asymptotic procedure is available. Denote by S_n^{X-Y} and S_n^{X+Y} the sample variances of $X_i - Y_i$, $i = 1, \dots, n$, and $X_i + Y_i$, $i = 1, \dots, n$. Let

$$B_n = \frac{S_n^{X+Y}}{S_n^{X-Y} + S_n^{X+Y}}.$$

Then $B_n = (W_n + 1)/2$ and

$$\sqrt{n-1} \frac{W_n}{\sqrt{1 - W_n^2}} = \sqrt{n-1} \frac{2B_n - 1}{2\sqrt{B_n(1 - B_n)}}.$$

To give the distribution of this test statistic in the case where $\rho = 0$, we can (and do) assume without loss of generality that the X_i, Y_i , $i = 1, \dots, n$, are independent unit normal variables. Then $(X_i - Y_i)/\sqrt{2}, (X_i + Y_i)/\sqrt{2}$, $i = 1, \dots, n$, are also independent unit normal variables. It follows that B_n has the Beta distribution $\mathcal{B}[(n-1)/2, (n-1)/2]$. This implies that

$$\sqrt{n-1} \frac{2B_n - 1}{2\sqrt{B_n(1 - B_n)}}$$

has the t_{n-1} distribution (Student's t distribution with $n-1$ degrees of freedom); see Devroye [6], p. 430. Thus, in the special case $d = 1$ we have a t -test for the testing problem $\mathcal{H} : \rho \leq 0$, $\mathcal{K} : \rho > 0$; the rejection region is

$$\left\{ \sqrt{n-1} \frac{W_n}{\sqrt{1 - W_n^2}} > t_{n-1;1-\alpha} \right\},$$

where, for given $\alpha \in (0, 1)$, $t_{n-1;1-\alpha}$ is the quantile of order $1 - \alpha$ of the t_{n-1} distribution. In fact, arguing as in Lehmann and Romano [16], Section 5.13, it can be shown that it is the uniformly most powerful unbiased test.

6. Numerical illustrations and simulations

The comparison approach used here is an asymptotic one. In order to check the extent to which these results are reflected in the power performance of the tests in finite sample cases, we consider an intraclass correlation model as a special case; see, e.g., Srivastava and Carter [27] for more details about this model.

Let $d = 2$ and let the covariance matrix Ξ defined in (2.2) be of the special form $\Xi = \sigma^2\{(1 - \rho)I_{2d} + \rho\mathbb{1}_{2d}\mathbb{1}_{2d}^\top\}$ where $\mathbb{1}_{2d} = (1, \dots, 1)^\top \in \mathbb{R}^{2d}$, $\sigma^2 > 0$ is the common variance of the components of the random d -vectors X_1 and Y_1 , and $\rho \in (-1/3, 1)$ is the intraclass correlation coefficient. As the power of the Hotelling tests does not depend on σ^2 , we choose $\sigma^2 = 1$. The mean vectors are assumed to be $a = E(X_1) = 0$ and $\mu = b = E(Y_1) = (0.4, 0.4)^\top$. Due to the fact that the test statistics T_n and T'_n have noncentral F distributions, power calculations are easily done. For given level $\alpha = 0.05$ and intraclass correlation coefficient chosen to be $\rho = 1/6$, $\rho = -1/6$, and $\rho = 1/20$, Figure 1 shows the power curves of the tests as functions of the sample size $n \in \{5, \dots, 100\}$. A positive (negative) parameter value of ρ corresponds to the hypothesis $\mathcal{H}_1$, meaning that the Hotelling test in paired samples dominates its counterpart in independent samples ($\mathcal{H}_2$, meaning that the Hotelling test in independent samples dominates its counterpart in paired samples). For positive (negative) values of ρ , one expects to see that the power curve of the Hotelling test in paired samples is above (below) the power curve of the Hotelling test in independent samples. In fact, we find that this is realized, at least if the sample size n is large enough (precisely, $n \geq 7$ if $\rho = 1/6$, $n \geq 18$ if $\rho = 1/20$).

To give a first insight into the power performance of the tests presented in Section 5, we provide the result of a small simulation study. It deals with the testing problem (5.1) and the test associated with $T_{1,n}$ (the $T_{1,n}$ test) only. The special model considered is most easily described by the distributional identities

$$X_1 \stackrel{\mathcal{D}}{=} U_1 + \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} W_1, \quad Y_1 \stackrel{\mathcal{D}}{=} V_1 + \epsilon \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} W_1$$

with parameters $\epsilon \in \{-1, +1\}$, $0 \leq \rho_i < \infty$, $i = 1, 2$, $\rho_1 \neq \rho_2$, and independent bivariate random vectors U_1, V_1, W_1 with the same distribution $\mathcal{N}_2(0, I_2)$. So, again $d = 2$.

The submatrices Π and Σ are given by

$$\Pi = \epsilon \begin{pmatrix} \rho_1^2 & 0 \\ 0 & \rho_2^2 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 1+\rho_1^2 & 0 \\ 0 & 1+\rho_2^2 \end{pmatrix}.$$

One has $p = (\epsilon\rho_1^2, 0, 0, \epsilon\rho_2^2)^\top$ and $\ell_1(p) = \min(\epsilon\rho_1^2, \epsilon\rho_2^2)$. The parameter value $\epsilon = 1$ ($\epsilon = -1$) corresponds to the hypothesis $\mathcal{H}_1$ (alternative $\mathcal{K}_1$). If $\epsilon = 1$, it is useful to distinguish the cases $0 = \rho_1^2 < \rho_2^2$, say, and $\rho_1^2 > 0, \rho_2^2 > 0$.

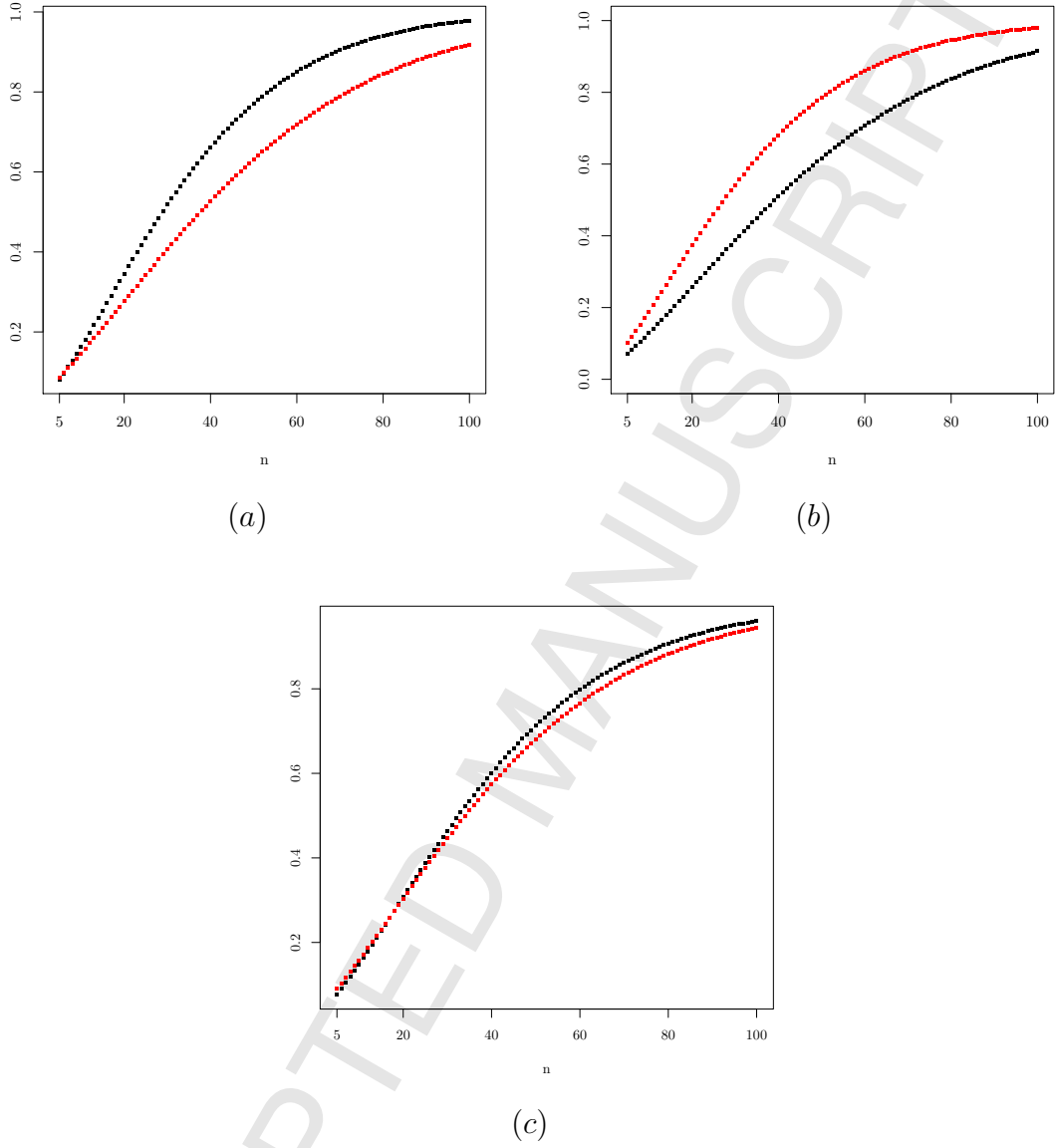


Figure 1: Power curves of the Hotelling tests in paired samples (black) and independent samples (red); sample size $5 \leq n \leq 100$; intraclass correlation parameter $\rho = 1/6$ (a); $\rho = -1/6$ (b); $\rho = 1/20$ (c)

If $\epsilon = 1$ and $\rho_1^2 = 0, \rho_2^2 > 0$ then $\ell_1(p) = 0$; we then speak of an underlying distribution belonging to the “boundary” of the hypothesis $\mathcal{H}_1$. If $\epsilon = 1$ and $\rho_1^2 > 0, \rho_2^2 > 0$ then $\ell_1(p) > 0$, in which case we speak of an underlying distribution belonging to the “interior” of the hypothesis $\mathcal{H}_1$.

By Theorem 5.1, for large n the error probability of the $T_{1,n}$ test should be approximately

Table 1: Power of the $T_{1,n}$ test; significance level $\alpha = 0.05$. The first (second) row in each line shows the power for the sample size $n = 100$ ($n = 200$).

ϵ	ρ_1^2	ρ_2^2						
		1/4	1/2	3/4	1	4/3	2	4
+1	0	0.07012	0.06286	0.06310	0.05971	0.05766	0.05600	0.05475
		0.06725	0.05858	0.05703	0.05690	0.05611	0.05392	0.05373
	1/100	0.05893	0.05035	0.04839	0.04681	0.04521	0.04614	0.04409
		0.04975	0.04553	0.04335	0.04087	0.04092	0.04058	0.03804
	9/100	0.01328	0.00968	0.00923	0.00932	0.00855	0.00816	0.00784
		0.00482	0.00355	0.00320	0.00284	0.00279	0.00295	0.00292
		ρ_2^2						
		1/50	1/30	1/20	1/15	1/8	1/5	1/3
-1	0	0.20883	0.22939	0.26494	0.30283	0.44343	0.63256	0.86868
		0.22749	0.26517	0.32074	0.38116	0.61046	0.84171	0.98456
	1/100	0.22261	0.24535	0.28096	0.31662	0.45758	0.64079	0.87353
		0.24802	0.28605	0.33964	0.39664	0.62163	0.84533	0.98501
	9/100	0.38453	0.40437	0.43515	0.46057	0.57648	0.72599	0.90446
		0.50717	0.53206	0.57251	0.61384	0.75922	0.90592	0.99039

equal to the given level α in the first case; it should be approximately equal to zero in the second case. Table 1 shows the power of the $T_{1,n}$ test for given level $\alpha = 0.05$, sample size $n \in \{100, 200\}$ and parameter values $\rho_1^2 \in \{0, 1/100, 9/100\}$ and $\rho_2^2 \in \{1/4, 1/2, 3/4, 1, 4/3, 2, 4\}$ if $\epsilon = +1$, $\rho_2^2 \in \{1/50, 1/30, 1/20, 1/15, 1/8, 1/5, 1/3\}$ if $\epsilon = -1$. The findings are in accordance with the theoretical considerations.

7. Conclusion and outlook

For given parameter values Π and Σ , Theorem 4.1 provides a concise answer to the question that in terms of classical asymptotic efficiency concepts one of the Hotelling tests in paired samples or independent samples dominates the other one. It would be of interest to obtain similar results on the basis of other efficiency notions. For the Chernoff efficiency, for example, which is similar in nature to the Hodges–Lehmann efficiency, it should be possible to present an answer in the same vein.

Section 5, which dealt with pretests for superiority, is by no means complete. Of course,

there are other testing problems of interest, e.g., the testing problem

$$\begin{aligned}\mathcal{H}_5 : \quad & \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} \geq 1 \text{ and } \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} \leq 1, \\ \mathcal{K}_5 : \quad & \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} < 1 \text{ or } \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} \frac{\mu^\top (2\Sigma)^{-1} \mu}{\mu^\top \Lambda^{-1} \mu} > 1.\end{aligned}$$

For a first approach to this testing problem one may be led to use test statistics based on the extreme eigenvalues $\ell_{1,n}$ and $\ell_{d,n}$, the test statistic $T_{5,n} = \max(\sqrt{n}\ell_{1,n}, -\sqrt{n}\ell_{d,n})$, for example. We postpone the exploration of the necessary theoretical details to be done with this or other testing problems to future work.

Acknowledgements

The authors thank the Editor, Christian Genest, and the referees for constructive comments and suggestions. The second author was supported by a doctoral scholarship from the Hans-Böckler-Stiftung.

References

- [1] Anderson, T. W., 2003. *An introduction to multivariate statistical analysis*. Wiley, New York.
- [2] Arcones, M. A., 2005. Bahadur efficiency of the likelihood ratio test. *Math. Methods Statist.* 14, 163–179.
- [3] Bahadur, R. R., 1967. An optimal property of the likelihood ratio statistic. *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability, Vol. I, Statistics*, pp. 13–26.
- [4] Bahadur, R. R., 1971. *Some limit theorems in statistics*. SIAM, Philadelphia.
- [5] Baringhaus, L., 1987. Asymptotic optimality of multivariate linear hypothesis tests. *J. Multivariate Anal.* 23, 303–311.
- [6] Devroye, L., 1986. *Non-uniform random variate generation*. Springer, New York.
- [7] Frischmuth, M., 1991. *Zweistichproben-Test oder Test bei verbundenen Stichproben?* Diploma thesis, Leibniz Universität Hannover.
- [8] Groeneboom, P. and Oosterhoff, J., 1977. Bahadur efficiency and probabilities of large deviations. *Statistica Neerlandica* 31, 1–24.
- [9] Gruber, M. H. J., 2013. *Matrix algebra for linear models*. Wiley, New York.
- [10] Kallenberg, W. C. M. and Kourouklis, S., 1992. Hodges–Lehmann optimality of tests. *Statist. Probab. Lett.* 14, 31–38.
- [11] Knüsel, L., 1977. Vergleich zweier Testverfahren zur Symmetriepfung in 4-Felder-Tafeln. *Metrika* 24, 175–186.
- [12] Königsberger, K., 2004. *Analysis 2*. Springer, Berlin.
- [13] Kourouklis, S., 1988. Hodges–Lehmann-eficiencies of certain tests in multivariate analysis and regression analysis. *Canad. J. Statist.* 16, 87–95.
- [14] Kourouklis, S., 1989. On the relation between Hodges–Lehmann efficiency and Pitman efficiency. *Canad. J. Statist.* 17, 311–318.
- [15] Koziol, J. A., 1978. Exact slopes of certain multivariate tests of hypotheses. *Ann. Statist.* 6, 546–558.
- [16] Lehmann, E. and Romano, J., 2005. *Testing statistical hypotheses*. Wiley, New York.
- [17] Magnus, J. R., 1985. On differentiating eigenvalues and eigenvectors. *Econometric Theory* 1, 179–191.

- [18] Magnus, J. R. and Neudecker, H., 1988. *Matrix differential calculus with applications in statistics and econometrics*. Wiley, New York.
- [19] Muirhead, R. J., 2005. *Aspects of multivariate statistical theory*. Wiley, New York.
- [20] Nikitin, Y., 1995. *Asymptotic efficiency of nonparametric tests*. Cambridge University Press, Cambridge.
- [21] Pollak, M. and Cohen, J., 1981. A comparison of the independent-samples t-test and the paired-samples t-test when the observations are nonnegatively correlated pairs. *J. Statist. Plann. Inference* 5, 133–146.
- [22] Pratt, J., 1961. Length of confidence intervals. *J. Amer. Statist. Assoc.* 56, 549–567.
- [23] Puri, M. L. and Sen, P. K., 1971. *Nonparametric methods in multivariate analysis*. Wiley, New York.
- [24] Roverato, A. and Whittaker, J., 1998. The Isserlis matrix and its application to non-decomposable graphical Gaussian models. *Biometrika* 85, 3, 711–725.
- [25] Rublík, F., 1994. On Hodges–Lehmann optimality of LR tests. *Kybernetika*. 30, 199–210.
- [26] Serfling, R. S., 1980. *Approximation theorems of mathematical statistics*. Wiley, New York.
- [27] Srivastava, M.S. and Carter, E.M., 1983. *Applied multivariate statistics*. North-Holland, New York.
- [28] Titchmarsh, E. C., 1978. *The theory of functions*. Oxford University Press, Oxford.
- [29] Wieand, H. S., 1976. A Condition under which the Pitman and Bahadur approaches to efficiency coincide. *Ann. Statist.* 4, 1003–1011.
- [30] Wilkinson, J. H., 1965. *The algebraic eigenvalue problem*. Clarendon Press, Oxford.