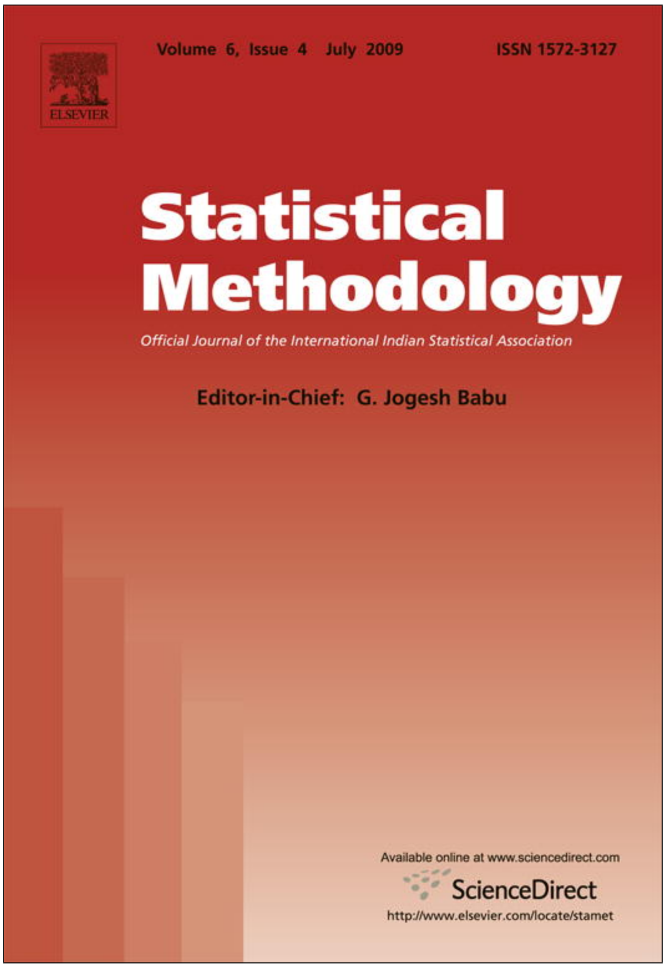
Volume 6, Issue 4 July 2009
ISSN 1572-3127
ELSEVIER
Statistical Methodology
Official Journal of the International Indian Statistical Association
Editor-in-Chief: G. Jogesh Babu
Available online at www.sciencedirect.com
ScienceDirect
http://www.elsevier.com/locate/stamet

Statistical Methodology 6 (2009) 424–436

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Rényi statistics for testing equality of autocorrelation coefficients[☆]

Tomáš Hobza [a,b,∗], Domingo Morales [c], Leandro Pardo [d]

[a] Department of Mathematics, Czech Technical University, Prague, Czech Republic
[b] Institute of Information Theory and Automation of the ASCR, Prague, Czech Republic
[c] Operation Research Center, Miguel Hernández University of Elche, Elche, Spain
[d] Department of Statistics, Complutense University of Madrid, Madrid, Spain

ARTICLE INFO

Article history:
Received 29 November 2007
Received in revised form
12 March 2009
Accepted 21 March 2009

Keywords:
Rényi divergence
Likelihood divergence statistics
Normal multivariate distribution
Testing hypothesis
Autocorrelation coefficient

ABSTRACT

The problem of testing for equality of autocorrelation coefficients of two populations in multivariate data when errors are autocorrelated is considered. We derive Rényi statistics defined as divergences between unrestricted and restricted estimated joint probability density functions and we show that they are asymptotically chi-square distributed under the null hypothesis of interest. Monte Carlo simulation experiments are carried out to investigate the behavior of Rényi statistics and to make comparisons with test statistics based on the approach of Bhandary [M. Bhandary, Test for equality of autocorrelation coefficients for two populations in multivariate data when the errors are autocorrelated, Statistics & Probability Letters 73 (2005) 333–342] for the problem under consideration. Rényi statistics showed to have significantly better behavior.

1. Introduction

In stationary time series models first order autocorrelation coefficients, ρ, measure the correlation between observations in two arbitrary instants $t-1$ and t. AR(1) processes (with $|\rho| < 1$) are an example of this type of stochastic processes, which is widely used to model time series of economic indicators. Statistical inference concerning ρ in one-sample problems has been extensively studied and many statistics have been obtained (see for instance [1–10] and the references therein),

☆ Supported by the grants MTM2006-05693, MSMT 1M0572, MTM 2006-06872 and UCM2006-910707.

∗ Corresponding address: Department of Mathematics, Czech Technical University of Prague, Trojanova 13, 120 00 Prague 2, Czech Republic. Tel.: +420 224 358 547; fax: +420 224 918 643.

E-mail address: hobza@fjfi.cvut.cz (T. Hobza).

1572-3127/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.stamet.2009.03.001

however the extension to two-sample problems has not received the same attention. We only know the test proposed by Bhandary [11]. This author considers the problem of testing the equality of autocorrelation coefficients of two independent multinormal samples with AR(1) structure. In this paper new statistics based on the Statistical Information Theory approach are introduced to treat the same problem.

For testing a general hypothesis about parameters of one population, the likelihood ratio test statistics is of general use. The likelihood ratio test statistic is a measure of deviation between the maximum likelihood achieved under the null hypothesis and the maximum achieved over the whole parameter space. Following this philosophy, a different measure of deviation, like a divergence, can be used. Some tests based on divergences have already been proposed in the literature, and it has been shown that in many cases they represent good competitors to classical tests. The history of the minimum divergence statistical method is so extensive that it deserves a review paper itself. Here we just would like to mention the concrete origins of the particular minimum divergence method for continuous data used in this paper.

Our first available reference on this issue is [12], who suggested to test a simple null hypothesis using the Kullback–Leibler divergence [13], providing its asymptotic distribution. Salicrú et al. [14] and Morales et al. [15] extended these results to the problem of testing composite hypotheses, using some families of divergences, like Csiszár's ϕ-divergence [16] or (h, ϕ)-divergences [17]. Some generalizations to dependent data and to multi-sample problems are given by Morales et al. [18,19] and Hobza et al. [20,21]. The first group of authors established a theoretical framework to apply divergence-based methods to testing simple and composite hypotheses in general exponential families. The second group concentrated its effort in testing hypotheses involving the parameters of s population without assuming any exponential model. As both approaches are equivalent, the test statistics appearing in this paper can be derived from any of the two. The first one makes a particularization from the general exponential family to the normal multivariate distribution. The second one requires similar calculations, as those appearing in [11], for the normal multivariate distribution and this is the main reason to choose this alternative to write the paper.

A well-known subfamily of (h, ϕ)-divergences is the Rényi family, which was introduced by Rényi [22] and defined in the nowadays form by Liese and Vajda [23]. Among the test statistics obtained from the Rényi divergences one can find the Kullback–Leibler statistic. An advantage of the Rényi family is that spelled-out formulas of the divergences can be obtained for some commonly used probability distributions. Some interesting books dealing with statistical applications of divergence measures are [23–26].

In this paper, we treat the problem of testing the equality of two autocorrelation coefficients based on two independent multivariate normal samples by means of statistics based on the Rényi family of divergences. More formally, let us consider the sample $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1}$ of n_1 i.i.d. observations from population 1 where

$$\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip}) \sim N_p(\boldsymbol{\mu}_1, \Sigma_1), \quad i = 1, \ldots, n_1.$$

Here N_p denotes the p-variate normal distribution with the mean vector $\boldsymbol{\mu}_1$ and the covariance matrix Σ_1 specified as $\boldsymbol{\mu}_1 = (\mu_{11}, \ldots \mu_{1p})$ and $\Sigma_1 = \Sigma(\sigma_1^2, \varrho_1)$, where the matrix Σ has the form

$$\Sigma = \Sigma(\sigma^2, \varrho) = \sigma^2 \begin{bmatrix} 1 & \varrho & \varrho^2 & \ldots & \varrho^{p-1} \\ \varrho & 1 & \varrho & \ldots & \varrho^{p-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \varrho^{p-1} & \varrho^{p-2} & \ldots & \ldots & 1 \end{bmatrix}. \tag{1.1}$$

In the matrix Σ_1, σ_1^2 represents the variance of each component and ϱ is called the autocorrelation coefficient. By $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2}$ we denote the sample of n_2 i.i.d. observations from population 2 where

$$\boldsymbol{y}_j = (y_{j1}, \ldots, y_{jp}) \sim N_p(\boldsymbol{\mu}_2, \Sigma_2), \quad j = 1, \ldots, n_2$$

and $\boldsymbol{\mu}_2 = (\mu_{21}, \ldots, \mu_{2p})$, $\Sigma_2 = \Sigma(\sigma_2^2, \varrho_2)$. Samples $\boldsymbol{x}_i$ (similarly $\boldsymbol{y}_j$) can be considered as realizations of an AR(1) time series model with the structure

$$X_t = \mu_{1t} + \rho_1(X_{t-1} - \mu_{1t-1}) + u_t, \quad \text{where } u_t \text{ i.i.d. } N(0, \sigma_1^2(1 - \rho_1^2)).$$

We are interested in testing the hypothesis

$$H_0 : \varrho_1 = \varrho_2 \quad \text{versus } H_1 : \varrho_1 \neq \varrho_2 \tag{1.2}$$

under the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$. We treat this problem in two different ways. In Section 2 we apply the approach of Bhandary [11] and derive formulas of the Rényi statistics based on a suitable transformation of data. The test statistic proposed by Bhandary is presented too. In Section 3 a "classical" approach is considered, i.e. we deal with the original non-transformed model and derive formulas to calculate the H_0-restricted and non-restricted MLEs. Further, formulas of the Rényi test statistics for this case are presented and on the basis of results in [21] their asymptotic distribution is given. In Section 4 we carry out simulation studies to determine the critical values for statistics based on the Bhandary's approach, to investigate a small sample behavior of the classical Rényi statistics and to compare it with the performance of the Bhandary-based statistics.

2. Rényi test statistics based on Bhandary's approach

To simplify the problem we follow the steps of [11] and do a suitable transformation of the data using the matrix

$$T = T(\varrho) = \begin{bmatrix} \sqrt{1-\varrho^2} & 0 & 0 & \dots & 0 \\ -\varrho & 1 & 0 & \dots & 0 \\ 0 & -\varrho & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -\varrho & 1 \end{bmatrix}. \tag{2.1}$$

The reason for such transformation is that $T\Sigma T' = \sigma^2(1-\varrho^2)I_p$, where I_p denotes identity matrix of order $p \times p$, and so we can obtain a model with diagonal covariance matrix. In this paper, the simplified notation $T_1 = T(\varrho_1)$ and $T_2 = T(\varrho_2)$ is used.

The first transformed sample, corresponding to population 1, is denoted by $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n_1}$ with

$$\boldsymbol{u}_i' = T_1\boldsymbol{x}_i', \qquad \boldsymbol{u}_i = (u_{i1}, \dots, u_{ip}) \sim N_p(\boldsymbol{\mu}_1^*, \Sigma_1^*),$$

$i = 1, \dots, n_1$, and the transformed mean vector and covariance matrix are $\boldsymbol{\mu}_1^* = T_1\boldsymbol{\mu}_1'$ and $\Sigma_1^* = \eta_1 I_p$, where $\eta_1 = \sigma^2(1-\varrho_1^2)$. Similarly, the transformed sample from population 2 is denoted by $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n_2}$ with

$$\boldsymbol{v}_j' = T_2\boldsymbol{y}_j', \qquad \boldsymbol{v}_j = (v_{j1}, \dots, v_{jp}) \sim N_p(\boldsymbol{\mu}_2^*, \Sigma_2^*),$$

$j = 1, \dots, n_2$, and $\boldsymbol{\mu}_2^* = T_2\boldsymbol{\mu}_2'$ and $\Sigma_2^* = \eta_2 I_p$, where $\eta_2 = \sigma^2(1-\varrho_2^2)$.

First we derive formulas for non-restricted and H_0-restricted MLEs of the parameters $\boldsymbol{\mu}_1^*, \boldsymbol{\mu}_2^*, \eta_1$ and η_2 of the transformed model. Maximizing the log-likelihood function based on the joint sample $\boldsymbol{w} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n_1}, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{n_2})$ with respect to $\boldsymbol{\mu}_1^*, \boldsymbol{\mu}_2^*, \eta_1, \eta_2$ we get the non-restricted maximum likelihood estimates

$$\widehat{\boldsymbol{\mu}}_1^* = \frac{1}{n_1}\sum_{i=1}^{n_1}\boldsymbol{u}_i \triangleq \overline{\boldsymbol{u}}, \qquad \widehat{\boldsymbol{\mu}}_2^* = \frac{1}{n_2}\sum_{j=1}^{n_2}\boldsymbol{v}_j \triangleq \overline{\boldsymbol{v}} \tag{2.2}$$

and

$$\widehat{\eta}_1 = \frac{1}{n_1 p}\sum_{i=1}^{n_1}(\boldsymbol{u}_i - \overline{\boldsymbol{u}})(\boldsymbol{u}_i - \overline{\boldsymbol{u}})', \qquad \widehat{\eta}_2 = \frac{1}{n_2 p}\sum_{j=1}^{n_2}(\boldsymbol{v}_j - \overline{\boldsymbol{v}})(\boldsymbol{v}_j - \overline{\boldsymbol{v}})'. \tag{2.3}$$

Similarly, maximizing the log-likelihood function with respect to $\boldsymbol{\mu}_1^*, \boldsymbol{\mu}_2^*$ and η under $H_0 : \varrho_1 = \varrho_2$, where $\eta = \sigma^2(1-\varrho^2)$ and $\varrho = \varrho_1 = \varrho_2$, we get the restricted maximum likelihood estimates

$$\widehat{\boldsymbol{\mu}}_1^{*0} = \overline{\boldsymbol{u}}, \qquad \widehat{\boldsymbol{\mu}}_2^{*0} = \overline{\boldsymbol{v}} \quad \text{and} \quad \widehat{\eta} = \frac{1}{n}(n_1\widehat{\eta}_1 + n_2\widehat{\eta}_2), \tag{2.4}$$

where $n = n_1 + n_2$ and $\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}, \widehat{\eta}_1$ and $\widehat{\eta}_2$ are defined in (2.2) and (2.3).

The Rényi divergence between two multivariate normal distributions $N_d(\boldsymbol{\nu}_1, \boldsymbol{\Delta}_1)$ and $N_d(\boldsymbol{\nu}_0, \boldsymbol{\Delta}_0)$ is (see e.g. formula (3.3) on page 174 in [27] where it is derived for $\alpha \in [0, 1]$ or Proposition 2.22 on page 43 in [23] where it is obtained more generally for exponential families and $\alpha \in \mathbb{R}$)

$$D_a\left((\boldsymbol{\nu}_1, \boldsymbol{\Delta}_1), (\boldsymbol{\nu}_0, \boldsymbol{\Delta}_0)\right) = \frac{1}{2}\left[(\boldsymbol{\nu}_1 - \boldsymbol{\nu}_0)(a\boldsymbol{\Delta}_0 + (1-a)\boldsymbol{\Delta}_1)^{-1}(\boldsymbol{\nu}_1 - \boldsymbol{\nu}_0)'\right] - \frac{1}{2}\,\frac{1}{a(a-1)}\log\frac{|a\boldsymbol{\Delta}_0 + (1-a)\boldsymbol{\Delta}_1|}{|\boldsymbol{\Delta}_1|^{1-a}|\boldsymbol{\Delta}_0|^{a}}, \tag{2.5}$$

$a \in \mathbb{R} - \{0, 1\}$. We are interested in the Rényi divergence between the non-restricted and the restricted likelihood of the joint sample $\boldsymbol{w}$ and thus we have, respectively

$$\boldsymbol{\nu}_1 = (\overline{\boldsymbol{u}}, \ldots, \overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}, \ldots, \overline{\boldsymbol{v}})_{1\times d}, \qquad \boldsymbol{\Delta}_1 = \operatorname{diag}(\widehat{\eta}_1, \ldots, \widehat{\eta}_1, \widehat{\eta}_2, \ldots, \widehat{\eta}_2)_{d\times d} \tag{2.6}$$

and

$$\boldsymbol{\nu}_0 = \boldsymbol{\nu}_1, \qquad \boldsymbol{\Delta}_0 = \operatorname{diag}(\widehat{\eta}, \widehat{\eta}, \ldots, \widehat{\eta})_{d\times d} \tag{2.7}$$

for $d = (n_1 + n_2)p = np$. Substituting these expressions in (2.5), we get the Rényi divergence between the likelihoods

$$D_a(\widehat{\eta}_1, \widehat{\eta}_2, \widehat{\eta}) \triangleq D_a\left((\boldsymbol{\nu}_1, \boldsymbol{\Delta}_1), (\boldsymbol{\nu}_0, \boldsymbol{\Delta}_0)\right) = -\frac{1}{2}\,\frac{1}{a(a-1)}[n_1 p \log(a\widehat{\eta} + (1-a)\widehat{\eta}_1) + n_2 p \log(a\widehat{\eta} + (1-a)\widehat{\eta}_2) - (1-a)p[n_1 \log\widehat{\eta}_1 + n_2 \log\widehat{\eta}_2] - anp\log\widehat{\eta}]. \tag{2.8}$$

Using the transformation matrices T_1, T_2 and the original samples, the estimates $\widehat{\eta}_1, \widehat{\eta}_2, \widehat{\eta}$ can be written as

$$\begin{aligned}
\widehat{\eta}_1(\varrho_1) &= \frac{1}{n_1 p}\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})T_1'T_1(\boldsymbol{x}_i - \overline{\boldsymbol{x}})' \\
&= \frac{1}{n_1 p}\left[\sum_{i=1}^{n_1}\left(\sum_{r=1}^{p}(x_{ir} - \overline{x}_r)^2 - 2\varrho_1\sum_{r=2}^{p}(x_{ir} - \overline{x}_r)(x_{i,r-1} - \overline{x}_{r-1}) + \varrho_1^2\sum_{r=2}^{p-1}(x_{ir} - \overline{x}_r)^2\right)\right], \\
\widehat{\eta}_2(\varrho_2) &= \frac{1}{n_2 p}\sum_{j=1}^{n_2}(\boldsymbol{y}_j - \overline{\boldsymbol{y}})T_2'T_2(\boldsymbol{y}_j - \overline{\boldsymbol{y}})' \\
&= \frac{1}{n_2 p}\left[\sum_{j=1}^{n_2}\left(\sum_{s=1}^{p}(y_{js} - \overline{y}_s)^2 - 2\varrho_2\sum_{s=2}^{p}(y_{js} - \overline{y}_s)(y_{j,s-1} - \overline{y}_{s-1}) + \varrho_2^2\sum_{s=2}^{p-1}(y_{js} - \overline{y}_s)^2\right)\right]
\end{aligned}$$

and

$$\widehat{\eta}(\varrho_1, \varrho_2) = \frac{1}{n}(n_1\widehat{\eta}_1(\varrho_1) + n_2\widehat{\eta}_2(\varrho_2)).$$

Since the above expressions depend on the unknown parameters ϱ_1, ϱ_2 we now follow the steps of [11]. Under the null hypotheses it holds $\varrho_1 = \varrho_2$ and thus we define the test statistics R_a as

$$R_a = \max_{-1\le\varrho\le 1} D_a(\widehat{\eta}_1(\varrho), \widehat{\eta}_2(\varrho), \widehat{\eta}(\varrho, \varrho)), \tag{2.9}$$

where $D_a(\widehat{\eta}_1, \widehat{\eta}_2, \widehat{\eta})$ is defined in (2.8). The corresponding Rényi divergence α-test is

$$\text{reject } H_0 \quad \text{if } R_a > R_{a,1-\alpha}, \tag{2.10}$$

where the critical value $R_{a,1-\alpha}$, giving a right tail probability of α, may be determined by simulation.

The Kullback–Leibler divergence can be obtained as a limit of the Rényi divergence D_a for $a \to 1$ and will be denoted by D_1. Its formula for two multivariate normal distributions $N(\boldsymbol{\nu}_1, \boldsymbol{\Delta}_1)$ and $N(\boldsymbol{\nu}_0, \boldsymbol{\Delta}_0)$ is (see e.g. [27] or [23])

$$D_1\left((\boldsymbol{\nu}_1, \boldsymbol{\Delta}_1), (\boldsymbol{\nu}_0, \boldsymbol{\Delta}_0)\right) = \frac{1}{2}\left[(\boldsymbol{\nu}_1 - \boldsymbol{\nu}_0)\boldsymbol{\Delta}_0^{-1}(\boldsymbol{\nu}_1 - \boldsymbol{\nu}_0)' + \operatorname{trace}\left(\boldsymbol{\Delta}_0^{-1}\boldsymbol{\Delta}_1 - I\right) + \log\frac{|\boldsymbol{\Delta}_0|}{|\boldsymbol{\Delta}_1|}\right]. \tag{2.11}$$

Plugging into this formula the expressions (2.6) and (2.7) we get the Kullback–Leibler divergence between likelihoods

$$D_1(\widehat{\eta}_1, \widehat{\eta}_2, \widehat{\eta}) \triangleq D_1\left((\boldsymbol{\nu}_1, \boldsymbol{\Delta}_1), (\boldsymbol{\nu}_0, \boldsymbol{\Delta}_0)\right) = \frac{1}{2}(np \log \widehat{\eta} - n_1 p \log \widehat{\eta}_1 - n_2 p \log \widehat{\eta}_2). \tag{2.12}$$

Let us note that $D_1(\widehat{\eta}_1, \widehat{\eta}_2, \widehat{\eta}) = -\log \Lambda$, where Λ is the likelihood ratio test statistics obtained on page 338 of [11]. Similarly as in the previous case, the corresponding Kullback–Leibler divergence α-test of H_0 is defined by (2.9) and (2.10) for $a = 1$.

Under the present setup [11] derived a test statistic of the form

$$\Lambda^* = \lambda_{\max}(AB^{-1}), \tag{2.13}$$

where $\lambda_{\max}(D)$ is the largest eigenvalue of the matrix D and A, B are matrices from $\mathbb{R}^{2,2}$ of the form

$$A = \begin{bmatrix} \sum_{i=1}^{n_1}\sum_{r=1}^{p}(x_{ir} - \bar{x}_r)^2 & -\sum_{i=1}^{n_1}\sum_{r=2}^{p}(x_{ir} - \bar{x}_r)(x_{i,r-1} - \bar{x}_{r-1}) \\ -\sum_{i=1}^{n_1}\sum_{r=2}^{p}(x_{ir} - \bar{x}_r)(x_{i,r-1} - \bar{x}_{r-1}) & \sum_{i=1}^{n_1}\sum_{r=2}^{p-1}(x_{ir} - \bar{x}_r)^2 \end{bmatrix},$$

$$B = \begin{bmatrix} \sum_{j=1}^{n_2}\sum_{s=1}^{p}(y_{js} - \bar{y}_s)^2 & -\sum_{j=1}^{n_2}\sum_{s=2}^{p}(y_{js} - \bar{y}_s)(y_{j,s-1} - \bar{y}_{s-1}) \\ -\sum_{j=1}^{n_2}\sum_{s=2}^{p}(y_{js} - \bar{y}_s)(y_{j,s-1} - \bar{y}_{s-1}) & \sum_{j=1}^{n_2}\sum_{s=2}^{p-1}(y_{js} - \bar{y}_s)^2 \end{bmatrix}.$$

The corresponding Bhandary test of H_0 is

$$\text{reject } H_0 \quad \text{if} \quad \lambda_{\max}(AB^{-1}) > \lambda_{1-\alpha},$$

where the critical value $\lambda_{1-\alpha}$ has again to be determined by simulation.

3. Classical approach

In this section we treat the problem of testing H_0 by means of the Rényi divergence without the simplifying transformation of the model used by Bhandary. Consider the log-likelihood function based on the original joint sample $\boldsymbol{z} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2})$, which is given by

$$\ell = -\frac{np}{2}\log 2\pi - \frac{n_1}{2}\log|\Sigma_1| - \frac{n_2}{2}\log|\Sigma_2| - \frac{1}{2}\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)\Sigma_1^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)' - \frac{1}{2}\sum_{j=1}^{n_2}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)\Sigma_2^{-1}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)'. \tag{3.1}$$

We describe a procedure to obtain MLEs of parameters under the hypothesis $H_0 : \varrho_1 = \varrho_2 = \varrho$ when log-likelihood function to be maximized, ℓ_0, is obtained from (3.1) with $\Sigma_1 = \Sigma_2 = \Sigma$. Equating to zero the first derivatives of ℓ_0 with respect to $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ one gets

$$\widehat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{x}} = \frac{1}{n_1}\sum_{i=1}^{n_1}\boldsymbol{x}_i \quad \text{and} \quad \widehat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{y}} = \frac{1}{n_2}\sum_{j=1}^{n_2}\boldsymbol{y}_j. \tag{3.2}$$

The first derivatives of ℓ_0 with respect to σ^2 and ϱ are

$$S_{\sigma^2} = \frac{\partial \ell_0}{\partial \sigma^2} = -\frac{np}{2\sigma^2} + \frac{1}{2\sigma^2}\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)\Sigma^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)' + \frac{1}{2\sigma^2}\sum_{j=1}^{n_2}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)\Sigma^{-1}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)' \tag{3.3}$$

and

$$S_\varrho = \frac{\partial \ell_0}{\partial \varrho} = \frac{n(p-1)\varrho}{1-\varrho^2} + \frac{1}{2}\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)\Sigma^{-1}\frac{\partial \Sigma}{\partial \varrho}\Sigma^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)' + \frac{1}{2}\sum_{j=1}^{n_2}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)\Sigma^{-1}\frac{\partial \Sigma}{\partial \varrho}\Sigma^{-1}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)'. \tag{3.4}$$

The maximum likelihood estimates of σ^2 and ϱ can be computed numerically using the Fisher-scoring algorithm for which we need the Fisher information matrix

$$F = \begin{pmatrix} F_{\sigma^2\sigma^2} & F_{\sigma^2\varrho} \\ F_{\varrho\sigma^2} & F_{\varrho\varrho} \end{pmatrix}. \tag{3.5}$$

Calculating the expectations of the second order derivatives of the log-likelihood function ℓ_0 we get

$$F_{\sigma^2\sigma^2} = \frac{np}{2\sigma^4}, \qquad F_{\sigma^2\varrho} = -\frac{n(p-1)\varrho}{\sigma^2(1-\varrho^2)} \quad \text{and} \quad F_{\varrho\varrho} = \frac{n(p-1)(1+\varrho^2)}{(1-\varrho^2)^2}.$$

The iterative algorithm for computing the MLEs is then given by the equation

$$\theta^{\ell+1} = \theta^\ell + F^{-1}(\theta^\ell)S(\theta^\ell), \tag{3.6}$$

where $\theta = \left(\sigma^2, \varrho\right)'$, $S = \left(S_{\sigma^2}, S_\varrho\right)'$. To initiate the algorithm we take the seeds

$$\sigma_0^2 = \frac{1}{p(n-2)}\left\{\sum_{i=1}^{n_1}\sum_{r=1}^{p}(x_{ir} - \bar{x}_r)^2 + \sum_{j=1}^{n_2}\sum_{s=1}^{p}(y_{js} - \bar{y}_s)^2\right\}, \quad \varrho_0 = 0 \tag{3.7}$$

for

$$\bar{x}_r = \frac{1}{n_1}\sum_{i=1}^{n_1} x_{ir} \quad \text{and} \quad \bar{y}_s = \frac{1}{n_2}\sum_{j=1}^{n_2} y_{js}.$$

In the case of no restriction on the parameters we get the same estimates $\widehat{\boldsymbol{\mu}}_1$ and $\widehat{\boldsymbol{\mu}}_2$ as in (3.2) and the first derivatives of the log-likelihood function ℓ given in (3.1) with respect to the remaining parameters σ^2, ϱ_1 and ϱ_2 are

$$S_{\sigma^2} = \frac{\partial l}{\partial \sigma^2} = -\frac{np}{2\sigma^2} + \frac{1}{2\sigma^2}\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)\Sigma_1^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)' + \frac{1}{2\sigma^2}\sum_{j=1}^{n_2}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)\Sigma_2^{-1}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)',$$

$$S_{\varrho_1} = \frac{\partial l}{\partial \varrho_1} = \frac{n_1(p-1)\varrho_1}{1-\varrho_1^2} + \frac{1}{2}\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)\Sigma_1^{-1}\frac{\partial \Sigma_1}{\partial \varrho_1}\Sigma_1^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)'$$

and

$$S_{\varrho_2} = \frac{\partial l}{\partial \varrho_2} = \frac{n_2(p-1)\varrho_2}{1-\varrho_2^2} + \frac{1}{2}\sum_{j=1}^{n_2}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)\Sigma_2^{-1}\frac{\partial \Sigma_2}{\partial \varrho_2}\Sigma_2^{-1}(\boldsymbol{y}_j - \boldsymbol{\mu}_2)'.$$

The corresponding Fisher information matrix has the form

$$F = \begin{pmatrix} F_{\sigma^2\sigma^2} & F_{\sigma^2\varrho_1} & F_{\sigma^2\varrho_2} \\ F_{\varrho_1\sigma^2} & F_{\varrho_1\varrho_1} & F_{\varrho_1\varrho_2} \\ F_{\varrho_2\sigma^2} & F_{\varrho_2\varrho_1} & F_{\varrho_2\varrho_2} \end{pmatrix}$$

with

$$F_{\sigma^2\sigma^2} = \frac{np}{2\sigma^4}, \qquad F_{\sigma^2\varrho_k} = -\frac{n_k(p-1)\varrho_k}{\sigma^2(1-\varrho_k^2)}, \quad k = 1, 2, \qquad F_{\varrho_1\varrho_2} = 0$$

and

$$F_{\varrho_k\varrho_k} = \frac{n_k(p-1)(1+\varrho_k^2)}{(1-\varrho_k^2)^2}, \quad k = 1, 2.$$

In this case the updating equation of the Fisher-scoring algorithm is

$$\theta^{\ell+1} = \theta^{\ell} + F^{-1}(\theta^{\ell})S(\theta^{\ell}), \tag{3.8}$$

where $\theta = \left(\sigma^2, \varrho_1, \varrho_2\right)'$ and $S = \left(S_\sigma^2, S_{\varrho_1}, S_{\varrho_2}\right)'$. In the simulation experiment we take the seeds $\varrho_{10} = \varrho_{20} = 0$ and σ_0^2 the same as in (3.7).

Let us denote by $\widehat{\sigma}^2, \widehat{\varrho}_1$ and $\widehat{\varrho}_2$ the unrestricted MLEs of σ^2, ϱ_1 and ϱ_2 and by $\widehat{\sigma}_0^2, \widehat{\varrho}_0$ the H_0-restricted MLEs of the corresponding parameters σ^2, ϱ. Let us also define $d = (n_1 + n_2)p = np$ and

$$\widehat{\Sigma}_0 = \Sigma(\widehat{\sigma}_0^2, \widehat{\varrho}_0), \qquad \widehat{\Sigma}_1 = \Sigma(\widehat{\sigma}^2, \widehat{\varrho}_1), \qquad \widehat{\Sigma}_2 = \Sigma(\widehat{\sigma}^2, \widehat{\varrho}_2)$$

for matrix Σ introduced in (1.1). By substituting

$$\mathbf{\Delta_1} = \operatorname{diag}\left(\widehat{\Sigma}_1, \ldots, \widehat{\Sigma}_1, \widehat{\Sigma}_2, \ldots, \widehat{\Sigma}_2\right)_{d\times d} \quad \text{and} \quad \mathbf{\Delta_0} = \operatorname{diag}\left(\widehat{\Sigma}_0, \ldots, \widehat{\Sigma}_0\right)_{d\times d}$$

into the formula (2.5) (notice that $\boldsymbol{\nu}_1 = \boldsymbol{\nu}_0$) we get the Rényi divergence between the unrestricted and the H_0-restricted likelihoods of the joint sample $\boldsymbol{z}$

$$D_a = D_a\left(\widehat{\Sigma}_1, \widehat{\Sigma}_2, \widehat{\Sigma}_0\right) = -\frac{1}{2}\frac{1}{a(a-1)}\Big\{n_1 \log\left|a\widehat{\Sigma}_0 + (1-a)\widehat{\Sigma}_1\right| + n_2 \log\left|a\widehat{\Sigma}_0 + (1-a)\widehat{\Sigma}_2\right| \\ - (1-a)\left[n_1 \log\left|\widehat{\Sigma}_1\right| + n_2 \log\left|\widehat{\Sigma}_2\right|\right] - an \log\left|\widehat{\Sigma}_0\right|\Big\}.$$

Proposition 3.1. *Under the present model and the assumption that n_1, n_2 go to infinity at the same rate, i.e. there exist $\lambda_1, \lambda_2 \in (0, 1), \lambda_1 + \lambda_2 = 1$ such that $n_1/(n_1 + n_2) \to \lambda_1$ and $n_2/(n_1 + n_2) \to \lambda_2$ for $n_1 \to \infty, n_2 \to \infty$, the test statistics $2D_a$ is asymptotically χ^2 distributed with one degree of freedom, i.e.*

$$2D_a \xrightarrow{L} \chi_1^2 \quad \text{as } n_1 \to \infty, n_2 \to \infty.$$

Proof. The proof follows directly by checking the assumptions (A1)–(A2) and (H1)–(H3) of Theorem 1 in [21]. These assumptions basically deals with the integrability of the multivariate normal density of original joint sample $\boldsymbol{z}$. This has been done by applying straightforward and cumbersome calculations, which are not presented here. □

The proposed "classical" Rényi test of H_0 with asymptotic significance level α is

$$\text{reject } H_0 \quad \text{if } R_a^C \triangleq 2\,D_a > \chi^2_{1,1-\alpha}. \tag{3.9}$$

Using the formula (2.11) we get the Kullback–Leibler divergence between likelihoods of the joint sample $\boldsymbol{z}$ in the form

$$D_1 = D_1\left(\widehat{\Sigma}_1, \widehat{\Sigma}_2, \widehat{\Sigma}_0\right) = \frac{1}{2}\Bigg[\frac{\widehat{\sigma}^2}{\widehat{\sigma}_0^2(1-\widehat{\varrho}_0^2)}\left(2n - p\widehat{\varrho}_0(n_1\widehat{\varrho}_1 + n_2\widehat{\varrho}_2) + n(p-2)(1+\widehat{\varrho}_0^2)\right) - np \\ + n\log\left|\widehat{\Sigma}_0\right| - n_1 \log\left|\widehat{\Sigma}_1\right| - n_2 \log\left|\widehat{\Sigma}_2\right|\Bigg].$$

Again, as in the case of the Rényi divergence, the corresponding test of asymptotic significance level α is

$$\text{reject } H_0 \quad \text{if } R_1^C \triangleq 2\,D_1 > \chi^2_{1,1-\alpha}.$$

Table 1
Estimated critical values $R_{a,0.95}$ and $\lambda_{0.95}$ of the test statistics R_a and Λ^* for testing H_0 (cf. (1.2)) at level $\alpha = 0.05$ with sample sizes $n_1 = 50, n_2 = 50$.

ϱ	$a = 0.5$	$a = 0.75$	$a = 1$	$a = 1.25$	$a = 1.5$	$a = 1.75$	$a = 2$	$a = 2.25$	Λ^*
0	4.635	4.590	4.558	4.539	4.532	4.538	4.558	4.591	1.540
0.1	4.540	4.496	4.466	4.447	4.441	4.447	4.466	4.498	1.528
0.2	4.718	4.671	4.638	4.618	4.611	4.618	4.638	4.673	1.532
0.3	4.788	4.740	4.706	4.685	4.678	4.685	4.706	4.741	1.539
0.4	5.064	5.010	4.972	4.949	4.942	4.949	4.972	5.012	1.562
0.5	5.205	5.148	5.108	5.084	5.076	5.084	5.108	5.151	1.567
0.6	5.645	5.578	5.531	5.503	5.493	5.502	5.531	5.581	1.597
0.7	5.981	5.907	5.854	5.822	5.811	5.822	5.854	5.910	1.630
0.8	6.724	6.630	6.564	6.524	6.511	6.523	6.564	6.635	1.718

4. Simulation experiment

To calculate the critical values, the sizes and powers of the considered statistics and to compare their small sample behavior we carry out simulation studies using the Matlab software. The dimension of the observation vectors is selected to be $p = 4$, the sample sizes are taken from the set $(n_1, n_2) \in \{(25, 25), (25, 50), (50, 50), (50, 100)\}$ and the mean vectors and variances are chosen as $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = (0, 0, 0, 0)$ and $\sigma_1 = \sigma_2 = 1$. The Bhandary statistic Λ^* (cf. (2.13)) and the Rényi statistics R_a (cf. (2.9)) and R_a^C (cf. (3.9)) for the values of a from the set $\mathcal{A} = \{0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25\}$ are considered and the level of the tests is set to be $\alpha = 0.05$.

To be able to use the Bhandary statistics Λ^* and the Rényi statistics R_a we need first to calculate the corresponding critical values. For this purpose we select the common values of the autocorrelation coefficients, $\rho = \rho_1 = \rho_2$, from the set $\Omega = \{0.1 \cdot i \mid i = 0, 1, \ldots, 8\}$. The critical values $\lambda_{0.95}$ and $R_{a,0.95}$ of the statistics Λ^* and R_a, $a \in \mathcal{A}$, are estimated as the 95th percentile of 10000 independent realizations of the statistics Λ^*, R_a obtained from samples generated for selected combination of parameters of the model. For illustration the calculated critical values for the case $n_1 = n_2 = 50$ are presented in the Table 1. More extensive study of the critical values can be found in [28].

We observe that the simulated critical values depend on the value of ϱ. In practice, to apply this approach it is necessary to: (a) estimate the autocorrelation coefficient from the joint sample, say $\widehat{\rho}$; (b) apply the above described algorithm under the assumption $\rho_1 = \rho_2 = \widehat{\rho}$; and (c) apply the test with the obtained critical value. From the algorithm to simulate the random samples it is further not difficult to see that the critical values do not depend on the value of the parameter $\sigma = \sigma_1 = \sigma_2$.

Once critical values are estimated, the estimates of powers of the given statistics are computed from the proportion of rejection of the null hypothesis using the estimated critical values.

For the study of powers two alternative lines to the null hypothesis H_0 are selected

$$\begin{aligned} &\text{(a)}\ \varrho_1 = 0.1 \quad \text{and} \quad \varrho_2 = 0.1 + 0.05i, \quad i \in \{-4, -3, \ldots, 0, \ldots, 4\}, \\ &\text{(b)}\ \varrho_1 = 0.5 \quad \text{and} \quad \varrho_2 = 0.5 + 0.05i, \quad i \in \{-4, -3, \ldots, 0, \ldots, 4\} \end{aligned} \tag{4.1}$$

and the following algorithm is used:

1. For the concrete choice of ϱ_1, ϱ_2 and n_1, n_2 generate samples $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1}$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2}$.
2. Calculate values of the Bhandary statistic Λ^* and of the Rényi statistics R_a for $a \in \mathcal{A} = \{0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25\}$. In the latter case the maximization over the region $\varrho \in (-1, 1)$ is done by using an algorithm based on the golden section search with parabolic interpolation (see [29,30]).
3. From the samples calculate the restricted maximum likelihood estimates $\widehat{\sigma}_0^2, \widehat{\varrho}_0$ of the parameters σ^2, ϱ and the non-restricted MLEs $\widehat{\sigma}^2, \widehat{\varrho}_1, \widehat{\varrho}_2$ of the parameters $\sigma^2, \varrho_1, \varrho_2$ using the Fisher-scoring algorithms (3.6) and (3.8), respectively.
4. Plugging these estimates into the corresponding formulas, the values of the test statistics R_a^C (cf. (3.9)), $a \in \mathcal{A}$, are obtained.

Table 2
Minima and maxima of the estimated sizes $\widehat{\alpha}_a^C$ of R_a^C for testing H_0 (cf. (1.2)) at the level $\alpha = 0.05$ based upon 10 000 samples of sizes n_1 and n_2.

n_1	n_2		$a = 0.5$	$a = 0.75$	$a = 1$	$a = 1.25$	$a = 1.5$	$a = 1.75$	$a = 2$	$a = 2.25$
25	25	min	0.0601	0.0590	0.0584	0.0579	0.0574	0.0579	0.0582	0.0593
		max	0.0730	0.0722	0.0721	0.0727	0.0737	0.0758	0.0778	0.0812
25	50	min	0.0589	0.0568	0.0556	0.0550	0.0537	0.0537	0.0539	0.0545
		max	0.0759	0.0750	0.0752	0.0754	0.0761	0.0774	0.0793	0.0827
50	50	min	0.0521	0.0513	0.0507	0.0506	0.0505	0.0505	0.0507	0.0519
		max	0.0609	0.0602	0.0602	0.0604	0.0606	0.0611	0.0621	0.0634
50	100	min	0.0525	0.0515	0.0506	0.0510	0.0513	0.0511	0.0515	0.0513
		max	0.0605	0.0604	0.0604	0.0608	0.0610	0.0613	0.0627	0.0640

5. The steps 1–4 are repeated 10^4 times and the estimated powers $\widehat{\beta}^*$, $\widehat{\beta}_a$ and $\widehat{\beta}_a^C$ of the considered test statistics Λ^*, R_a and R_a^C, $a \in \mathcal{A}$, are then calculated by the formulas
$$\widehat{\beta}^* = \frac{\#\{\Lambda^* > \lambda_{0.95}\}}{10^4}, \qquad \widehat{\beta}_a = \frac{\#\{R_a > R_{a,0.95}\}}{10^4} \quad \text{and} \quad \widehat{\beta}_a^C = \frac{\#\{R_a^C > \chi^2_{1,0.95}\}}{10^4},$$
where $\#\{condition\}$ denotes the number of repetitions for which *condition* is true. Note that the critical values $\lambda_{0.95}$, $R_{a,0.95}$ depend on the sample sizes n_1, n_2 and ϱ_1, ϱ_2.

In the case of the statistics R_a^C it is not necessary to evaluate the critical values but since we use the asymptotic distribution of the statistics before comparing the powers we should look at the sizes of the corresponding tests. The estimated sizes $\widehat{\alpha}_a^C$ of the Rényi statistics R_a^C are calculated by the steps 1,3,4,5 of the presented algorithm, the only difference is that samples are generated from models with $\varrho_1 = \varrho_2 = \varrho \in \Omega$. Minimal and maximal values
$$\min_{\varrho \in \Omega} \widehat{\alpha}_a^C \quad \text{and} \quad \max_{\varrho \in \Omega} \widehat{\alpha}_a^C$$
of the estimated sizes are presented in the Table 2.

From this table one can see that the asymptotic distribution used to approximate the distribution of our test statistics works quite well and for selected sample sizes the obtained estimated sizes are reasonably close to the desired level 0.05. The values of parameter a for which the estimated sizes are closer to 0.05 are most often those from the set $\{1, 1.25, 1.5\}$. Generally, the following trend can be also observed: with increasing a the estimated sizes decrease till some value of a, usually one of those from the above mentioned set, and then start to increase.

Let us now return to the study of powers. Since all the considered statistics R_a have almost the same behavior in the selected models and for the sake of brevity we present in the Tables 3–6 just the estimated powers for the statistics R_a^C and Λ^* and the sample sizes from the set $(n_1, n_2) \in \{(25, 25), (50, 100)\}$. Some of the estimated powers of the statistics R_a will be presented graphically at the end of this section. For more detailed description of the experiment and more complete results we refer to [28].

Concerning the powers of the statistics R_a^C the conclusion following from the presented tables are similar to those obtained from tables for the statistics R_a not presented here. For equal sample sizes the powers of all statistics R_a^C are very similar following the same trend as observed for sizes. Reasonable choice thus would be to select the statistics with the best size, i.e. with $a \in [1, 1.5]$. For unequal sample sizes and the case $\varrho = 0.1$ we observe increasing powers with increasing a. In the case $\varrho_1 = 0.5$ for $\varrho_2 < \varrho_1$ the powers decrease with increasing a and for $\varrho_2 > \varrho_1$ the powers increase with increasing a. Since the differences are not dramatic in some sense optimal choice would be again to select $a \in [1, 1.5]$.

Since behavior of the Rényi tests is in both supposed approaches very similar, for a visual comparison we select just one member of the family so that everything is more simple and transparent. The value $a = 1.25$ was chosen as a member of the above proposed set and so behaviors of the Rényi statistics $R_{1.25}$, the classical Rényi statistics $R_{1.25}^C$ and the Bhandary statistics Λ^* are compared.

Table 3
Estimated powers $\widehat{\beta}_a^C$ and $\widehat{\beta}^*$ of R_a^C and Λ^* for testing H_0 (cf. (1.2)) at the level $\alpha = 0.05$ based upon 10 000 samples of sizes $n_1 = 25, n_2 = 25$ and $\varrho_1 = 0.1$.

ϱ_2	$a = 0.5$	$a = 0.75$	$a = 1$	$a = 1.25$	$a = 1.5$	$a = 1.75$	$a = 2$	$a = 2.25$	Λ^*
−0.10	0.2557	0.2552	0.2553	0.2563	0.2590	0.2616	0.2666	0.2725	0.0992
−0.05	0.1726	0.1719	0.1720	0.1726	0.1749	0.1780	0.1827	0.1881	0.0698
0.00	0.1179	0.1172	0.1171	0.1180	0.1193	0.1210	0.1247	0.1288	0.0578
0.05	0.0818	0.0807	0.0809	0.0811	0.0828	0.0853	0.0883	0.0927	0.0478
0.10	0.0712	0.0703	0.0704	0.0708	0.0722	0.0743	0.0774	0.0807	0.0440
0.15	0.0813	0.0806	0.0805	0.0812	0.0825	0.0855	0.0884	0.0917	0.0522
0.20	0.1190	0.1186	0.1185	0.1196	0.1206	0.1229	0.1258	0.1297	0.0645
0.25	0.1889	0.1875	0.1874	0.1876	0.1897	0.1925	0.1970	0.2011	0.0901
0.30	0.2862	0.2844	0.2842	0.2857	0.2882	0.2930	0.2971	0.3043	0.1292

Table 4
Estimated powers $\widehat{\beta}_a^C$ and $\widehat{\beta}^*$ of R_a^C and Λ^* for testing H_0 (cf. (1.2)) at the level $\alpha = 0.05$ based upon 10 000 samples of sizes $n_1 = 25, n_2 = 25$ and $\varrho_1 = 0.5$.

ϱ_2	$a = 0.5$	$a = 0.75$	$a = 1$	$a = 1.25$	$a = 1.5$	$a = 1.75$	$a = 2$	$a = 2.25$	Λ^*
0.30	0.3724	0.3700	0.3691	0.3700	0.3718	0.3749	0.3794	0.3833	0.0657
0.35	0.2530	0.2512	0.2504	0.2502	0.2517	0.2536	0.2564	0.2602	0.0595
0.40	0.1552	0.1534	0.1526	0.1529	0.1534	0.1553	0.1572	0.1601	0.0473
0.45	0.0867	0.0862	0.0856	0.0852	0.0858	0.0865	0.0878	0.0900	0.0461
0.50	0.0622	0.0615	0.0609	0.0608	0.0614	0.0624	0.0632	0.0648	0.0503
0.55	0.0935	0.0910	0.0906	0.0903	0.0902	0.0909	0.0922	0.0945	0.0730
0.60	0.1874	0.1849	0.1841	0.1838	0.1838	0.1842	0.1864	0.1897	0.1240
0.65	0.3737	0.3715	0.3693	0.3685	0.3689	0.3704	0.3733	0.3781	0.2536
0.70	0.6290	0.6263	0.6238	0.6228	0.6238	0.6259	0.6282	0.6317	0.4748

Table 5
Estimated powers $\widehat{\beta}_a^C$ and $\widehat{\beta}^*$ of R_a^C and Λ^* for testing H_0 (cf. (1.2)) at the level $\alpha = 0.05$ based upon 10 000 samples of sizes $n_1 = 50, n_2 = 100$ and $\varrho_1 = 0.1$.

ϱ_2	$a = 0.5$	$a = 0.75$	$a = 1$	$a = 1.25$	$a = 1.5$	$a = 1.75$	$a = 2$	$a = 2.25$	Λ^*
−0.10	0.5292	0.5288	0.5287	0.5303	0.5327	0.5347	0.5370	0.5407	0.2132
−0.05	0.3432	0.3426	0.3421	0.3426	0.3441	0.3461	0.3482	0.3516	0.1391
0.00	0.1830	0.1822	0.1817	0.1820	0.1822	0.1834	0.1851	0.1867	0.0889
0.05	0.0907	0.0902	0.0898	0.0899	0.0900	0.0904	0.0909	0.0932	0.0579
0.10	0.0596	0.0593	0.0592	0.0594	0.0597	0.0600	0.0612	0.0624	0.0501
0.15	0.0928	0.0928	0.0931	0.0939	0.0949	0.0965	0.0983	0.1007	0.0664
0.20	0.1827	0.1836	0.1845	0.1862	0.1879	0.1901	0.1928	0.1955	0.1076
0.25	0.3487	0.3497	0.3511	0.3539	0.3563	0.3605	0.3646	0.3709	0.1868
0.30	0.5630	0.5647	0.5678	0.5703	0.5735	0.5778	0.5827	0.5890	0.3286

Table 6
Estimated powers $\widehat{\beta}_a^C$ and $\widehat{\beta}^*$ of R_a^C and Λ^* for testing H_0 (cf. (1.2)) at the level $\alpha = 0.05$ based upon 10 000 samples of sizes $n_1 = 50, n_2 = 100$ and $\varrho_1 = 0.5$.

ϱ_2	$a = 0.5$	$a = 0.75$	$a = 1$	$a = 1.25$	$a = 1.5$	$a = 1.75$	$a = 2$	$a = 2.25$	Λ^*
0.30	0.7521	0.7483	0.7449	0.7419	0.7400	0.7375	0.7362	0.7350	0.1247
0.35	0.5392	0.5344	0.5289	0.5253	0.5215	0.5178	0.5158	0.5136	0.0914
0.40	0.3079	0.3040	0.2994	0.2946	0.2913	0.2890	0.2871	0.2851	0.0668
0.45	0.1318	0.1283	0.1258	0.1227	0.1202	0.1184	0.1172	0.1164	0.0505
0.50	0.0561	0.0553	0.0551	0.0550	0.0549	0.0554	0.0554	0.0563	0.0512
0.55	0.1198	0.1221	0.1238	0.1258	0.1281	0.1315	0.1344	0.1377	0.1028
0.60	0.3605	0.3626	0.3671	0.3721	0.3770	0.3842	0.3908	0.3975	0.2758
0.65	0.7196	0.7222	0.7263	0.7306	0.7349	0.7396	0.7460	0.7515	0.6138
0.70	0.9499	0.9509	0.9520	0.9534	0.9544	0.9558	0.9572	0.9591	0.9087

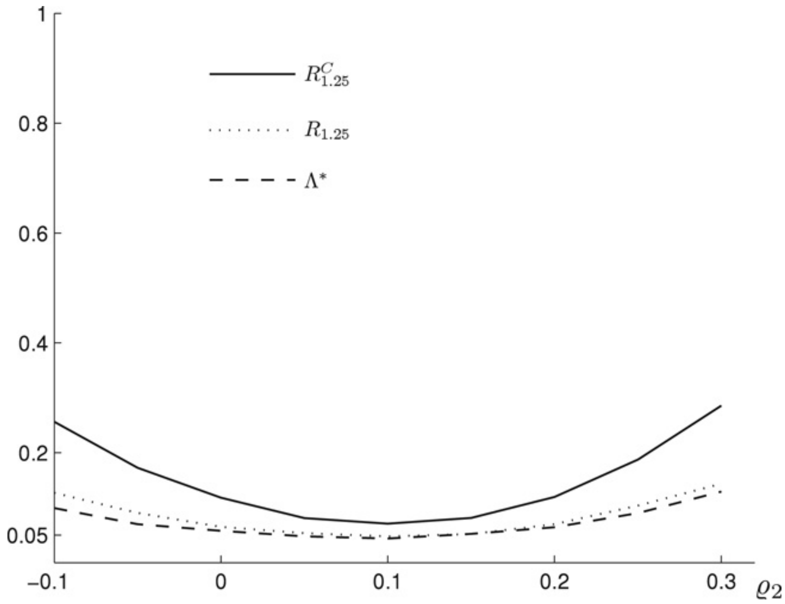

Fig. 1. Powers of the selected statistics for $n_1 = 25$, $n_2 = 25$, and $\varrho_1 = 0.1$.

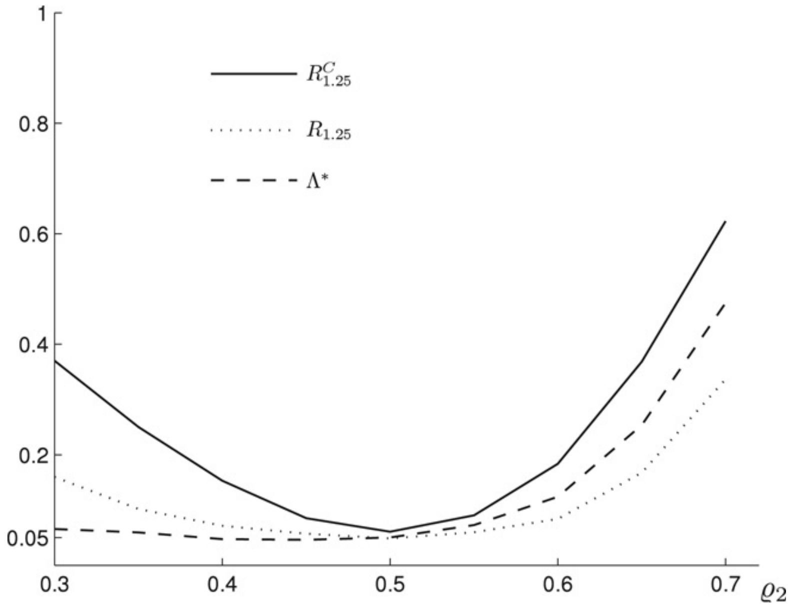

Fig. 2. Powers of the selected statistics for $n_1 = 25$, $n_2 = 25$, and $\varrho_1 = 0.5$

It is interesting to note that the value $a = 1.25$ was found in [31] in goodness-of-fit for multinomial populations as the value of "a" with higher order of convergence of the exact distribution to the asymptotic one.

Powers of the selected statistics are represented in Figs. 1–4. From these figures we deduce that the classical Rényi statistics $R^C_{1.25}$ has the best behavior in the sense of powers, for all presented situations. For the case $\varrho_1 = 0.1$ the Rényi statistics $R_{1.25}$ is slightly better than the Bhandary statistic Λ^* and for $\varrho_1 = 0.5$ it is better for $\varrho_2 < \varrho_1$ and slightly worse for $\varrho_2 > \varrho_1$.

In general, the Bhandary statistic Λ^* showed a poor behavior on the left hand side of the null hypothesis H_0 (i.e. $\varrho_2 < \varrho_1$) which is caused by its asymmetric definition (cf. (2.13)). Some attention should be thus paid to the order of the samples (i.e. which sample is sample 1) when using the Bhandary statistic.

As a conclusion we can say that the classical Rényi statistic R^C_a with $a \in [1, 1.5]$ can be recommended for use. The advantage of these statistics is that we need not simulate critical values

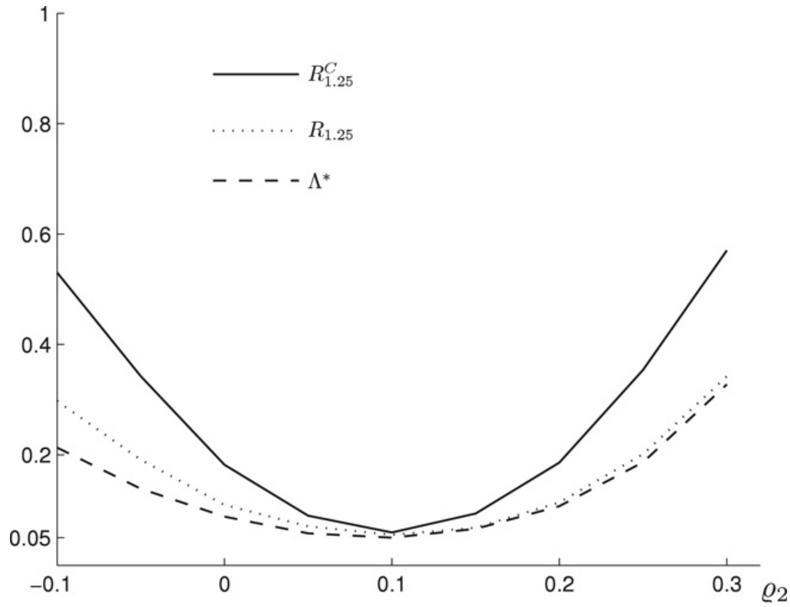

Fig. 3. Powers of the selected statistics for $n_1 = 50$, $n_2 = 100$, and $\varrho_1 = 0.1$.

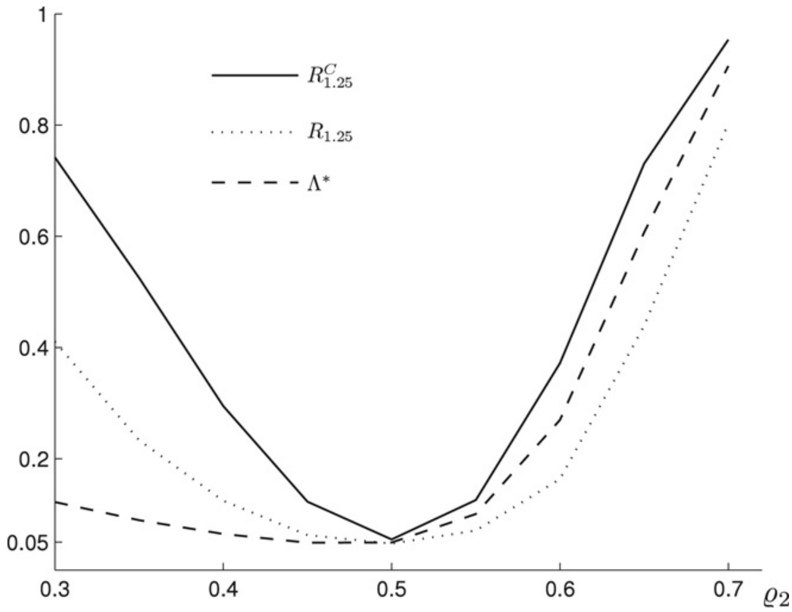

Fig. 4. Powers of the selected statistics for $n_1 = 50$, $n_2 = 100$, and $\varrho_1 = 0.5$.

and take care about the order of the samples as in the case of the Bhandary statistics and moreover they are expected to provide considerably higher powers. On the other hand the MLEs of the unknown parameters must be calculated numerically but the formulas are given and are simple and the Fisher-scoring algorithm converges very quickly in the supposed normal model.

Acknowledgements

We would like to thank the referees for their helpful comments. One of the authors is grateful to the Operation Research Center for its financial support during his stay at the Miguel Hernández University of Elche.

References

[1] J. Durbin, G.S. Watson, Testing serial correlation in least squares regression: I, Biometrika 37 (1950) 409–428.
[2] J. Durbin, G.S. Watson, Testing serial correlation in least squares regression: II, Biometrika 38 (1951) 159–178.

[3] J. Durbin, G.S. Watson, Testing serial correlation in least squares regression: III, Biometrika 58 (1971) 1–19.
[4] T.S. Breusch, Testing for autocorrelation in dynamic linear models, Australian Economic Papers 17 (1978) 334–355.
[5] L.G. Godfrey, Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables, Econometrica 46 (1978) 1293–1302.
[6] G.E.P. Box, David A. Pierce, Distribution of residual autocorrelations in autoregressive-integrated moving average time, Journal of the American Statistical Association 65 (1970) 1509–1526.
[7] D. Edgerton, G. Shukur, Testing autocorrelation in a system perspective, Econometric Review 18 (4) (1999) 343–386.
[8] G.M. Ljung, G.E.P. Box, On a measure of lack of fit in time-series models, Biometrika 66 (1978) 265–270.
[9] J. Jeong, S. Chung, Bootstrap tests for autocorrelation, Computational Statistics and Data Analysis 38 (2001) 49–69.
[10] A. Hatemi-J, Multivariate tests for autocorrelation in the stable and unstable VAR models, Economic Modeling 21 (2004) 661–683.
[11] M. Bhandary, Test for equality of autocorrelation coefficients for two populations in multivariate data when the errors are autocorrelated, Statistics & Probability Letters 73 (2005) 333–342.
[12] M. Kupperman, Further applications of information theory to multivariate analysis and statistical inference, Ph.D. Dissertation, The George Washington University, 1957.
[13] S. Kullback, Information Theory and Statistics, Wiley, New York, 1959.
[14] M. Salicrú, M.L. Menéndez, D. Morales, L. Pardo, On the applications of divergence type measures in testing statistical hypotheses, Journal of Multivariate Analysis 51 (1994) 372–391.
[15] D. Morales, L. Pardo, I. Vajda, Some new statistics for testing composite hypotheses in parametric models, Journal of Multivariate Analysis 62 (1) (1997) 137–168.
[16] I. Csiszár, Eine Informationstheoretische Ungleidung und ihre Anwendung auf den Bewis del Ergodizität on Markhoffschen Ketten, Publications of the Mathematical Institute, Hungarian Academy of Sciences 8 (1963) 84–108.
[17] M.L. Menéndez, D. Morales, L. Pardo, M. Salicrú, Asymptotic behavior and statistical applications of divergence measures in multinomial populations: A unified study, Statistical Papers 36 (1) (1995) 1–29.
[18] D. Morales, L. Pardo, I. Vajda, Rényi statistics in directed families of exponential experiments, Statistics 34 (2000) 151–174.
[19] D. Morales, L. Pardo, M.C. Pardo, I. Vajda, Rényi statistics for testing composite hypotheses in general exponential models, Statistics 38 (2) (2004) 133–147.
[20] T. Hobza, I. Molina, D. Morales, Likelihood divergence statistics for testing hypotheses in familial data, Communications in Statistics (Theory and Methods) 32 (2) (2003) 415–434.
[21] T. Hobza, I. Molina, D. Morales, Rényi statistics for testing hypotheses with s samples. Research report I-2007-30 in the Operations Research Center of the Miguel Hernández University of Elche. Available on the internet at: http://cio.umh.es/ES/publicaciones/ficheros/CIO_2007_30.pdf, 2007.
[22] A. Rényi, On measures of entropy and information, in: Fourth Berkeley Symposium on Mathematics, Statistics and Probability, vol. 1, University of California Press, Berkeley, 1961, pp. 547–561.
[23] F. Liese, I. Vajda, Convex Statistical Distances, Teubner, Leipzig, 1987.
[24] T.R.C. Read, N.A.C. Cressie, Goodness-of-fit Statistics for Discrete Multivariate Data, Springer-Verlag, New York, 1988.
[25] I. Vajda, Theory of Statistical Inference and Information, Kluwer Academic, Dordrecht, 1989.
[26] L. Pardo, Statistical Inference Based on Divergence Measures, Chapman & Hall/CRC, New York, 2006.
[27] J. Burbea, The convexity with respect to Gaussian distributions of divergences of order α, Utilitas Mathematica 26 (1984) 171–192.
[28] T. Hobza, D. Morales, L. Pardo, Testing equality of autocorrelation coefficients in multivariate normal models, Research report I-2007-15 in the Operations Research Center of the Miguel Hernández University of Elche. Available on the internet at: http://cio.umh.es/ES/publicaciones/ficheros/CIO_2007_15.pdf, 2007.
[29] G.E. Forsythe, M.A. Malcolm, C.B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall, 1976.
[30] Richard P. Brent, Algorithms for Minimization without Derivatives, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
[31] E. Landaburu, L. Pardo, Matching moments for a closer approximation of the weighted (h, ϕ)-divergence test statistics in goodness-of-fit for finite samples, Journal of the Franklin Institute 342 (2005) 115–129.