Testing equality of correlation coefficients in two populations via permutation methods.

M. Omelka[a], M. Pauly[b]

[a] *Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, 186 75 Praha 8*
[b] *Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, CH-3012 Bern*

Abstract

The present paper investigates the asymptotic behaviour of a studentized permutation test for testing equality of (Pearson) correlation coefficients in two populations. It is shown that this test is asymptotically of exact level and has the same power for contiguous alternatives as the corresponding asymptotic test. As a by-product we specify the assumptions needed for the validity of the permutation test suggested in Sakaori (2002). A small simulation study compares the finite sample properties of the considered tests.

Keywords: Behrens-Fisher problem, contiguous alternatives, permutation tests, Pearson correlation coefficient, power of a test, studentized statistics, two-sample tests.
2000 MSC: 62G09, 62G10

1. Introduction

Suppose we observe two independent random samples of sizes n_1 and n_2 of independent bivariate random vectors

$$\begin{aligned} X_1 &= (X_1^{(1)}, X_1^{(2)})^t, \ldots,\ X_{n_1} = (X_{n_1}^{(1)}, X_{n_1}^{(2)})^t \qquad \text{i.i.d. } F, \\ Y_1 &= (Y_1^{(1)}, Y_1^{(2)})^t, \ldots,\ Y_{n_2} = (Y_{n_2}^{(1)}, Y_{n_2}^{(2)})^t \qquad \text{i.i.d. } G, \end{aligned} \tag{1}$$

that are defined on a common probability space $(\Omega, \mathcal{A}, P)$ and put $n = n_1 + n_2$. Suppose that both covariance matrices of X_1 and Y_1 are positive definite and that

$$E(\|X_1\|^4 + \|Y_1\|^4) < \infty, \tag{2}$$

where $\|\cdot\|$ stands for an Euclidean norm. Put

$$\rho_1 = \frac{\operatorname{cov}(X_1^{(1)}, X_1^{(2)})}{\sqrt{\operatorname{var}(X_1^{(1)}) \operatorname{var}(X_1^{(2)})}} \quad \text{and} \quad \rho_2 = \frac{\operatorname{cov}(Y_1^{(1)}, Y_1^{(2)})}{\sqrt{\operatorname{var}(Y_1^{(1)}) \operatorname{var}(Y_1^{(2)})}}$$

for the (Pearson product-moment) correlation coefficients corresponding to the vectors X_1 and Y_1. In the following we will be interested in testing the null-hypothesis of equal correlation coefficients $H_0 : \rho_1 = \rho_2$ against the one-sided alternative $H_1 : \rho_1 > \rho_2$.

There is a large amount of literature on testing properties of correlation matrices, but most of the authors consider testing the equality of coefficients within a single correlation matrix. The asymptotic χ^2-test for the equality of an arbitrary number of independent population correlation coefficients in the bivariate case was already given by Pearson and Wilks (1933) (see p. 374) and the problem was further studied by Paul (1989). Kullback (1967) has generalized this test to correlation matrices. A different test aiming at comparing only two correlation matrices was proposed in Jennrich (1970). But in all those papers normality of observations is required.

Nevertheless, a simple asymptotic test that does not require the normality of observations follows for instance from formula (1.1) given in Jennrich (1970). This test can be rewritten for the bivariate case as follows. Put R_{n_1} and R_{n_2} for the empirical correlation coefficients of the first and second sample respectively, that is

$$R_{n_1} := \frac{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i^{(1)} - \overline{X}_{n_1}^{(1)})(X_i^{(2)} - \overline{X}_{n_1}^{(2)})}{\sqrt{\frac{1}{n_1}\sum_{j=1}^{n_1}(X_j^{(1)} - \overline{X}_{n_1}^{(1)})^2 \frac{1}{n_1}\sum_{k=1}^{n_1}(X_k^{(2)} - \overline{X}_{n_1}^{(2)})^2}} =: \frac{E_{n_1}}{D_{n_1}} \tag{3}$$

and similarly

$$R_{n_2} := \frac{\frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i^{(1)} - \overline{Y}_{n_2}^{(1)})(Y_i^{(2)} - \overline{Y}_{n_2}^{(2)})}{\sqrt{\frac{1}{n_2}\sum_{j=1}^{n_2}(Y_j^{(1)} - \overline{Y}_{n_2}^{(1)})^2 \frac{1}{n_2}\sum_{k=1}^{n_2}(Y_k^{(2)} - \overline{Y}_{n_1}^{(2)})^2}} =: \frac{E_{n_2}}{D_{n_2}}. \tag{4}$$

Then the test statistic is given by

$$\widetilde{T}_n := \frac{a_n(R_{n_1} - R_{n_2})}{V_n}, \tag{5}$$

where $a_n = \sqrt{\frac{n_1 n_2}{n}}$ is a normalizing sequence and V_n^2 is the estimate of the asymptotic variance of the difference $a_n(R_{n_1} - R_{n_2})$ that will be given later (see (14)).

In order to improve upon the small sample properties of the asymptotic test (5) Sakaori (2002) suggested a permutation test (φ_n^*) assuming that both samples follow the same distribution up to possible differences in locations and scales of the marginal distributions. His test is based on the difference $T_n = a_n(R_{n_1} - R_{n_2})$ and the significance of this statistic is assessed by the permutation of the *standardized pooled sample* $\mathbb{S}_n = (S_{n,1}, \ldots, S_{n,n})$, where

$$S_{n,i} := \left(\frac{X_i^{(j)} - \overline{X}_{n_1}^{(j)}}{\sqrt{\frac{1}{n_1} \sum_{k=1}^{n_1} (X_k^{(j)} - \overline{X}_{n_1}^{(j)})^2}} \right)_{j=1,2} \quad \text{for } 1 \leq i \leq n_1 \tag{6}$$

and

$$S_{n,n_1+i} := \left(\frac{Y_i^{(j)} - \overline{Y}_{n_2}^{(j)}}{\sqrt{\frac{1}{n_2} \sum_{k=1}^{n_2} (Y_k^{(j)} - \overline{Y}_{n_2}^{(j)})^2}} \right)_{j=1,2} \quad \text{for } 1 \leq i \leq n_2. \tag{7}$$

The validity of this test is only very briefly sketched in Sakaori (2002). In this paper the assumptions of this test are exactly specified. Furthermore, we suggest a permutation test that is based on the studentized statistic $\widetilde{T}_n$ given by (5). We will show that this test is asymptotically correct even if the bivariate distributions are different.

¿From the methodology point of view our paper follows the footsteps of Romano (1990), Neuhaus (1993) and Janssen (1997) who showed that under certain mild assumptions permutation tests are asymptotically correct even if the random variables are not exchangeable. More details on theory and applications of permutation tests can be found in the monographs of Edgington and Onghena (2007), Good (2005) and Pesarin and Salmaso (2010).

The paper is organized as follows. In Section 2, we describe the tests and state the asymptotic results. In Section 3, we investigate the finite sample properties of the tests. Section 4 adresses some possible extensions. The proofs are given in the Appendix.

2. Permutation tests for correlations

Recall that permutation tests are conditional tests given the data. In our situation the test can be formally constructed as follows. Let $\tau : (\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P}) \to \mathcal{S}_n$ be a random variable on some further probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P})$ that is uniformly distributed on the symmetric group $\mathcal{S}_n$ (the set of all permutations of $(1, \ldots, n)$) and independent of the data $X_1, \ldots, X_n$, where for simplicity of notations we set $X_{n_1+i} := Y_i$ for $1 \leq i \leq n_2$. Thereby independent means independence on the joint probability space $(\Omega \times \widetilde{\Omega}, \mathcal{A} \otimes \widetilde{\mathcal{A}}, P \otimes \widetilde{P})$, on which all random variables can be defined via projections.

Let $T_n = T_n((X_i)_{i \leq n})$ be a test statistic such that large values speak against the null hypothesis. Then for fixed observations $X_1(\omega), \ldots, X_n(\omega)$, $\omega \in \Omega$ the data dependent critical value $c_n(\alpha) = c_n(\alpha, \omega)$ is calculated as the $(1-\alpha)$-quantile of the permutation distribution of $T_n^* = T_n((X_{\tau(i)}(\omega))_{i \leq n})$.

2.1. Permutation test φ_n^*

First note that using the test statistic $T_n = a_n(R_{n_1} - R_{n_2})$ and permuting the original observations results in a test that is exact under a very restricted null hypothesis $\{F = G\}$, see e.g. Lehmann and Romano (2005). Unfortunately, the practical use of this test is very limited as it is not in general asymptotically correct if the distributions differ in any way (for instance only in location or scale). That is rather disappointing as a correlation coefficient is location and scale invariant. This leads naturally to the test φ_n^* suggested by Sakaori (2002), that permutes the *standardized pooled sample* $\mathbb{S}_n = (S_{n,1}, \ldots, S_{n,n})$ (defined by (6) and (7)).

To explore the asymptotic properties of the test φ_n^*, the following asymptotic representation of the statistic T_n is useful (for its derivation see the Appendix)

$$T_n - a_n(\rho_1 - \rho_2) = a_n \left(\frac{1}{n_1} \sum_{i=1}^{n_1} Z_i^X - \frac{1}{n_2} \sum_{i=1}^{n_2} Z_i^Y \right) + o_P(1), \tag{8}$$

where

$$\begin{aligned} Z_i^X &= \widetilde{X}_i^{(1)} \widetilde{X}_i^{(2)} - \tfrac{\rho_1}{2}\big((\widetilde{X}_i^{(1)})^2 + (\widetilde{X}_i^{(2)})^2\big), \quad 1 \leq i \leq n_1, & (9)\\ Z_i^Y &= \widetilde{Y}_i^{(1)} \widetilde{Y}_i^{(2)} - \tfrac{\rho_2}{2}\big((\widetilde{Y}_i^{(1)})^2 + (\widetilde{Y}_i^{(2)})^2\big), \quad 1 \leq i \leq n_2, & (10) \end{aligned}$$

with

$$\widetilde{X}_i^{(j)} = \frac{X_i^{(j)} - \mathrm{E}(X_i^{(j)})}{\sqrt{\mathrm{var}\,(X_i^{(j)})}} \quad \text{and} \quad \widetilde{Y}_i^{(j)} = \frac{Y_i^{(j)} - \mathrm{E}(Y_i^{(j)})}{\sqrt{\mathrm{var}\,(Y_i^{(j)})}}. \tag{11}$$

Further assume that $\frac{n_1}{n} \to p \in (0,1)$ as $n \to \infty$ and put $q = 1 - p$. With the help of (8) the central limit theorem yields that under the null hypothesis

$$T_n \xrightarrow{d} N(0, \sigma_Z^2), \qquad \text{where} \quad \sigma_Z^2 = q\,\sigma_X^2 + p\,\sigma_Y^2, \tag{12}$$

with $\sigma_X^2 = \mathrm{var}\,(Z_1^X)$ and $\sigma_Y^2 = \mathrm{var}\,(Z_1^Y)$. Here '$\xrightarrow{d}$' stands for convergence in distribution as $n \to \infty$ and $N(\mu, \sigma^2)$ for a normally distributed random variable with mean μ and variance σ^2.

Thus for the asymptotic validity of the test φ_n^* it is crucial that the permutation procedure asymptotically reproduces the null distribution of T_n. If so then one would expect that the permutation test φ_n^* has 'similar' asymptotic properties as the following test

$$\varphi_n = \mathbb{I}\{T_n > \sigma_Z\, u_{1-\alpha}\},$$

where $u_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard normal distribution $N(0,1)$. The asymptotic 'similarity' of the tests φ_n and φ_n^* is quantified by the following theorem.

Theorem 1. *Suppose that* (2) *and* $\frac{n_1}{n} \to p \in (0,1)$ *hold as* $n \to \infty$. *Further assume that either* $p = \frac{1}{2}$ *or* $\mathrm{var}\,(Z_1^X) = \mathrm{var}\,(Z_1^Y)$. *Then under the null hypothesis* H_0

$$E(|\varphi_n^* - \varphi_n|) \to 0 \qquad \text{as} \;\; n \to \infty. \tag{13}$$

A nice consequence of the L_1-convergence (13) is that φ_n^* and φ_n have not only the same asymptotic level, but also the same power functions for contiguous alternatives (e.g. see Section 6 of Janssen and Pauls (2003)).

Remark 1. Theorem 1 implies that the conjecture of Sakaori (2002) is right and the test φ_n^* is asymptotically valid provided the distributions F and G coincide up to differences in location and scale parameters as those do not affect the equation $\mathrm{var}\,(Z_1^X) = \mathrm{var}\,(Z_1^Y)$. Note that the validity also hold if the sample sizes are asymptotically balanced, that is $\frac{n_1}{n_2} \to 1$.

2.2. The studentized permutation test

The proof of Theorem 1 (given in Appendix) reveals that the validity of the test φ_n^* is not generally true as the permutation procedure interchanges the ratios p and q in the limit variance of the conditional permutation distribution. This is already a well-understood problem of permutation tests, see e.g. Romano (1990) or Janssen (1997). The solution to this problem is to permute the studentized test statistic $\widetilde{T}_n$ given by (5). The straightforward estimator of the variance of the statistic T_n is given by

$$V_n^2 = V_n^2(\mathbb{S}_n) := \frac{\hat{\sigma}_X^2}{n_1} + \frac{\hat{\sigma}_Y^2}{n_2}, \tag{14}$$

where

$$\hat{\sigma}_X^2 := \frac{1}{n_1-1}\sum_{i=1}^{n_1}(\hat{Z}_{n,i}-\overline{Z}_{n_1})^2, \quad \hat{\sigma}_Y^2 := \frac{1}{n_2-1}\sum_{i=n_1+1}^{n}(\hat{Z}_{n,i}-\overline{Z}_{n_2})^2$$

with

$$\hat{Z}_{n,i} := S_{n,i}^{(1)}S_{n,i}^{(2)} - \tfrac{1}{2}\Big(\tfrac{n_1}{n}R_{n_1} + \tfrac{n_2}{n}R_{n_2}\Big)\Big[(S_{n,i}^{(1)})^2 + (S_{n,i}^{(2)})^2\Big] \tag{15}$$

and

$$\overline{Z}_{n_1} := \frac{1}{n_1}\sum_{i=1}^{n_1}\hat{Z}_{n,i}, \quad \overline{Z}_{n_2} := \frac{1}{n_2}\sum_{i=n_1+1}^{n}\hat{Z}_{n,i}.$$

The studentized asymptotic test is defined as $\widetilde{\varphi}_n = \mathbb{I}\{\widetilde{T}_n > u_{1-\alpha}\}$, where $\widetilde{T}_n$ is defined in (5).

Put $\widetilde{\varphi}_n^*$ for the permutation version of the test $\widetilde{\varphi}_n$. Note that this test runs as follows. Let τ be a given permutation of the indices $\{1,\dots,n\}$. In the 'permutation world' the role of the original observations $(X_i)_i$ and $(Y_j)_j$ is played by the permuted pooled sample of standardized observations $\mathbb{S}_{n,\tau} = (S_{n,\tau(1)},\dots,S_{n,\tau(n)})$. Thus when computing $\widetilde{T}_n^*$ from this data one needs to standardize 'once more' and use

$$S_{n,\tau(i)}^* := \left(\frac{S_{n,\tau(i)}^{(j)} - \bar{S}_{n_1,\tau}^{(j)}}{\sqrt{\frac{1}{n_1}\sum_{k=1}^{n_1}(S_{n,\tau(k)}^{(j)} - \bar{S}_{n_1,\tau}^{(j)})^2}}\right)_{j=1,2}, \qquad 1 \le i \le n_1,$$

$$S_{n,\tau(n_1+i)}^* := \left(\frac{S_{n,\tau(n_1+i)}^{(j)} - \bar{S}_{n_2,\tau}^{(j)}}{\sqrt{\frac{1}{n_2}\sum_{k=1}^{n_2}(S_{n,\tau(n_1+k)}^{(j)} - \bar{S}_{n_2,\tau}^{(j)})^2}}\right)_{j=1,2}, \qquad 1 \le i \le n_2,$$

where

$$\bar{S}_{n_1,\tau}^{(j)} = \frac{1}{n_1}\sum_{k=1}^{n_1} S_{n,\tau(k)}^{(j)}, \qquad \bar{S}_{n_2,\tau}^{(j)} = \frac{1}{n_2}\sum_{k=n_1+1}^{n_1+n_2} S_{n,\tau(k)}^{(j)}.$$

Roughly speaking, in an analogous way as the original statistic $\widetilde{T}_n$ is a function of the standardized observation $\mathbb{S}_n$, the permuted statistic $\widetilde{T}_n^*$ is a function of $\mathbb{S}_{n,\tau}^* = (S_{n,\tau(1)}^*,\dots,S_{n,\tau(n)}^*)$. Formally, the permutation test is defined as $\widetilde{\varphi}_n^* = \mathbb{I}\{\widetilde{T}_n > \tilde{c}_n^*(\alpha)\}$, where $\tilde{c}_n^*(\alpha)$ is the conditional $(1-\alpha)$-quantile of $\widetilde{T}_n(\mathbb{S}_{n,\tau}^*)$.

Theorem 2. *Assume* (2) *and* $\frac{n_1}{n} \to p \in (0,1)$ *as* $n \to \infty$.

(a) If the null hypothesis holds, then

$$E(|\widetilde{\varphi}_n^* - \widetilde{\varphi}_n|) \to 0 \qquad as \quad n \to \infty.$$

(b) If $\rho_1 \neq \rho_2$, then for $n \to \infty$

$$E(\widetilde{\varphi}_n^*) \to \begin{cases} 1, & for\ \rho_1 > \rho_2, \\ 0, & for\ \rho_1 < \rho_2. \end{cases}$$

Theorem 2(a) implies that the permutation test $\widetilde{\varphi}_n^*$ is asymptotically of exact level under the null hypothesis. Moreover, the test $\widetilde{\varphi}_n^*$ has the same asymptotic power function for contiguous alternatives as the studentized asymptotic test $\widetilde{\varphi}_n$.

3. Simulations

A small Monte Carlo study was conducted to investigate the finite sample properties of the following tests:

- the asymptotic test $\widetilde{\varphi}_n = \mathbb{I}\{\widetilde{T}_n > u_{1-\alpha}\}$, where $\widetilde{T}_n$ is defined in (5),
- the (non-studentized) permutation test φ_n^* of Sakaori (2002) which is based on permuting the statistic $T_n = a_n(R_{n1} - R_{n2})$,
- the (studentized) permutation test $\widetilde{\varphi}_n^*$ which is based on permuting the statistic $\widetilde{T}_n$ (see Section 2.2).

The null hypothesis $H_0 : \rho_1 \leq \rho_2$ was tested against the one-sided alternative $H_1 : \rho_1 > \rho_2$ for the following model

$$X_j^{(2)} = \rho_1 X_j^{(1)} + \sqrt{1-\rho_1^2}\, e_j^X, \quad Y_j^{(2)} = \rho_2 Y_j^{(1)} + \sqrt{1-\rho_2^2}\, e_j^Y, \qquad (16)$$

where e_j^X is independent of $X_j^{(1)}$ and e_j^Y is independent of $Y_j^{(1)}$.

To illustrate our experience gained from the simulations we report only on the following scenarios

1. $X_j^{(1)}$, e_j^X as well as $Y_j^{(1)}$, e_j^Y follow the standard normal distribution $N(0,1)$.
2. $X_j^{(1)}$, e_j^X follow $N(0,1)$ and $Y_j^{(1)}$, e_j^Y follow a Laplace (double exponential) distribution with the density $f(x) = \frac{1}{2}\, e^{-|x|}$.

$\rho_1 = \rho_2$	-0.6	-0.3	0.0	0.3	0.6
	Scenario 1 (normal vs. normal).				
$\widetilde{\varphi}_n$	0.065	0.073	0.077	0.077	0.064
$\widetilde{\varphi}_n^*$	0.046	0.047	0.052	0.055	0.054
φ_n^*	0.040	0.046	0.048	0.049	0.035
	Scenario 2 (normal vs. laplace).				
$\widetilde{\varphi}_n$	0.074	0.077	0.068	0.070	0.065
$\widetilde{\varphi}_n^*$	0.050	0.050	0.045	0.052	0.058
φ_n^*	0.050	0.053	0.040	0.044	0.044
	Scenario 3 (exponential vs. normal).				
$\widetilde{\varphi}_n$	0.067	0.073	0.076	0.076	0.072
$\widetilde{\varphi}_n^*$	0.048	0.049	0.055	0.054	0.055
φ_n^*	0.042	0.044	0.047	0.041	0.029

Table 1: Levels of the tests φ_n, φ_n^* and $\widetilde{\varphi}_n^*$ when $n_1 = 20, n_2 = 10$.

3. $X_j^{(1)}$, e_j^X follow an exponential distribution with the density $f(x) = e^{-x}\,\mathbb{I}\{x > 0\}$ and $Y_j^{(1)}$, e_j^Y follow $N(0,1)$.

Note that in all scenarios $\mathrm{corr}(X_j^{(1)}, X_j^{(2)}) = \rho_1$ and $\mathrm{corr}(Y_j^{(1)}, Y_j^{(2)}) = \rho_2$.

While the first scenario represents the situation when multivariate distributions generating the two samples coincide, the multivariate distributions differ for the other two scenarios. The second scenario compares the correlation coefficients in two symmetric multivariate distributions. Finally, in the third scenario only one of the multivariate distributions is symmetric.

Unless stated otherwise the sample sizes are $n_1 = 20, n_2 = 10$. The p-value of the permutation test was approximated by generating 999 random permutations (see page 158 of Davison and Hinkley (1997)). The p-value was compared with the level $\alpha = 0.05$. 10 000 repetitions were used to approximate the level as well as the power of the test. The simulations were conducted with the help of R-computing environment, version 2.12.1 (see R Development Core Team (2010)).

The levels of the test for the sample sizes $n_1 = 20$ and $n_2 = 10$ are to be found in Table 1. One can see that the asymptotic test $\widetilde{\varphi}_n$ has considerable difficulties in keeping the prescribed level for small sample sizes. The stu-

	$\rho_1 = \rho_2 = 0.0$				$\rho_1 = \rho_2 = 0.6$			
n_1	20	40	200	2000	20	40	200	2000
n_2	10	20	100	1000	10	20	100	1000
Scenario 1 (normal vs. normal)								
$\widetilde{\varphi}_n$	0.077	0.066	0.049	0.053	0.064	0.056	0.050	0.050
$\widetilde{\varphi}_n^*$	0.052	0.055	0.045	0.052	0.054	0.052	0.051	0.050
φ_n^*	0.048	0.052	0.045	0.051	0.035	0.043	0.050	0.050
Scenario 2 (normal vs. laplace).								
$\widetilde{\varphi}_n$	0.068	0.062	0.052	0.049	0.065	0.060	0.052	0.049
$\widetilde{\varphi}_n^*$	0.045	0.053	0.053	0.049	0.058	0.060	0.056	0.049
φ_n^*	0.040	0.050	0.050	0.049	0.044	0.054	0.061	0.060
Scenario 3 (exponential vs. normal).								
$\widetilde{\varphi}_n$	0.076	0.059	0.050	0.052	0.072	0.072	0.055	0.052
$\widetilde{\varphi}_n^*$	0.055	0.043	0.040	0.049	0.055	0.055	0.047	0.050
φ_n^*	0.047	0.044	0.046	0.052	0.029	0.029	0.036	0.033

Table 2: Scenario 2 (normal vs. laplace) – levels of the tests.

dentized permutation test $\widetilde{\varphi}_n^*$ does a much better job in this aspect, but in some situations it is also slightly anti-conservative (with the actual level usually between 0.045 and 0.058). Finally, the (non-studentized) permutation test φ_n^* seems to be conservative for small sample sizes.

It is interesting to note that the conservativeness of the test φ_n^* holds even in situations where asymptotic calculations predict the test to be anti-conservative. For instance for Scenario 2 (normal vs. laplace) the asymptotic level of the test φ_n^* is exactly 0.050 for $\rho_1 = \rho_2 = 0$, but 0.063 for $\rho_1 = \rho_2 = 0.6$. The actual levels of the tests in these two situations for different sample sizes are to be found in Table 2. One can see that the asymptotic results about the level of the tests need sample sizes larger than 100 to kick in. Note also that the asymptotic test $\widetilde{\varphi}_n$ can be recommended for moderate sample sizes (> 100).

Finally, we also investigated the power of the tests. The results for Scenario 1 and 3 are to be found in Table 3. Findings for Scenario 2 are similar.

Note that for small samples the power of the asymptotic test $\widetilde{\varphi}_n$ cannot be strictly compared with the power of the two competitors as the asymptotic test is very anti-conservative (see Table 1). But it is interesting to see that the studentized permutation test $\widetilde{\varphi}_n^*$ does not loose much in terms of power in comparison with the asymptotic test, while having the levels

		Scenario 1 (normal vs normal)			Scenario 3 (exp. vs normal)		
ρ_1	ρ_2	$\widetilde{\varphi}_n$	$\widetilde{\varphi}_n^*$	φ_n^*	$\widetilde{\varphi}_n$	$\widetilde{\varphi}_n^*$	φ_n^*
0.9	0.6	0.508	0.530	0.369	0.462	0.443	0.213
0.9	0.3	0.796	0.818	0.690	0.736	0.718	0.484
0.9	0.0	0.927	0.940	0.885	0.887	0.878	0.717
0.9	-0.3	0.979	0.985	0.967	0.964	0.957	0.890
0.9	-0.6	0.997	0.998	0.996	0.993	0.992	0.980
0.9	-0.9	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.3	0.257	0.222	0.184	0.264	0.218	0.149
0.6	0.0	0.518	0.464	0.424	0.502	0.444	0.360
0.6	-0.3	0.773	0.722	0.706	0.743	0.690	0.635
0.6	-0.6	0.944	0.921	0.922	0.935	0.908	0.896
0.6	-0.9	0.999	0.998	0.998	0.998	0.998	0.998
0.3	0.0	0.220	0.170	0.160	0.232	0.186	0.160
0.3	-0.3	0.482	0.402	0.393	0.479	0.414	0.395
0.3	-0.6	0.786	0.708	0.720	0.812	0.757	0.759
0.3	-0.9	0.991	0.982	0.987	0.995	0.990	0.995
0.0	-0.3	0.219	0.162	0.159	0.230	0.178	0.172
0.0	-0.6	0.529	0.431	0.442	0.562	0.489	0.500
0.0	-0.9	0.957	0.922	0.931	0.974	0.960	0.970
-0.3	-0.6	0.267	0.196	0.196	0.267	0.213	0.203
-0.3	-0.9	0.843	0.779	0.769	0.848	0.804	0.792
-0.6	-0.9	0.556	0.485	0.425	0.477	0.425	0.370

Table 3: Power comparison of φ_n, φ_n^* and $\widetilde{\varphi}_n^*$ when $n_1 = 20, n_2 = 10$.

much closer to the prescribed value of 0.05. Similarly, as the test $\widetilde{\varphi}_n^*$ keeps approximately the level and the (non-studentized) permutation test φ_n^* is conservative in our scenarios, the power of $\widetilde{\varphi}_n^*$ is usually higher.

Our simulation experience may be summarized as follows.

- The asymptotic test seems to be the best option for samples larger than 100. For smaller samples the test is often quite anti-conservative.
- The non-studentized permutation test φ_n^* is usually conservative for small samples and thus it can be of interest in such situations. However, the price to pay can sometimes be a considerable lack of power for alternatives close to the null hypothesis. Note also that this is purely simulation experience for small samples that lacks theoretical justification. With moderate and large samples, the asymptotic results start to kick in and the test requires the assumptions of Theorem 1

to have asymptotically the correct level (which are e.g. not fulfilled in our Scenarios 2 and 3 if $\rho_1\rho_2 \neq 0$).

- The actual level of the studentized permutation test $\widetilde{\varphi}_n^*$ is usually between 0.045 and 0.058. If one is willing to accept this fact then the test $\widetilde{\varphi}_n^*$ is a good option for small samples.

4. Further discussion

For the brevity of the presentation we considered only the test statistic $a_n(R_{n1} - R_{n2})$ (as well as its studentized version (5)) and one-sided tests when investigating permutation tests for testing equality of correlation coefficients. Besides considering two-sided tests our results can be generalized in a straightforward way in the following directions.

4.1. Fisher's z-transformation

Inspired by the bivariate normal model one can try to stabilize the variance with the help of Fisher's z-transformation (see for instance Example 3.6 in Van der Vaart (2000)). That is why Sakaori (2002) suggested also a permutation test based on the following statistic

$$T_n^F = a_n \left[\tfrac{1}{2} \log\left(\tfrac{1+R_{n1}}{1-R_{n1}}\right) - \tfrac{1}{2} \log\left(\tfrac{1+R_{n2}}{1-R_{n2}}\right) \right].$$

By similar arguments as in the proof of (8) one can show that provided $\rho_1 = \rho_2$

$$T_n^F = \frac{a_n(R_2 - R_1)}{1 - \rho_1^2} + o_P(1).$$

Thus the asymptotic analysis of the permutation tests based on T_n^F can be derived by modifying the arguments given in our paper. Regarding the finite sample properties our simulation experience (not presented in this paper) is that the properties of the tests based on T_n^F are slightly better for normally distributed data but can be considerable worse in other situations.

4.2. Comparing multiple samples

Our results can also be generalized to the problem of testing the equality of correlation coefficients of k-independent samples. Let R_{ni} be the empirical correlation coefficient in the i-th sample and n_i the corresponding sample size. Provided that it is reasonable to assume that all the samples come from the same distribution (up to possible changes in scale and location in marginals), or provided that $\frac{n_i}{n} \approx \frac{1}{k}$ for $i = 1, \ldots, k$, one can use e.g. the test statistic

$$\sum_{i=1}^{k} n_i (R_{ni} - \overline{R}_n)^2, \qquad \text{where} \quad \overline{R}_n = \frac{1}{n} \sum_{i=1}^{k} n_i R_{ni}.$$

If neither of the above assumptions is satisfied then one can use one of the test statistics that are modifying ANOVA in presence of heteroscedasticity, see e.g. Argaç (2003) and the references therein.

4.3. Comparing more than one aspect of the data

In this paper we concentrated on comparison of Pearson correlation coefficients of two bivariate populations that may differ in all other aspects. Researches are very often interested in comparing the populations from more than one aspect. The general methodology for permutational approach to multiaspect tests can be found in Pesarin and Salmaso (2010) and a nice application to comparison of two multivariate distributions in Brombin et al. (2011). Note that with the help of the methodology described in those work, one can for instance construct tests of the joint equality of several measures of associations (e.g. Pearson correlations coefficient, Spearman's rho, Kendall's tau, ...) or the tests for equality of correlation matrices in more than two dimensions.

Appendix

Note that in the following, one can assume without loss of generality that the original observations $X_j^{(1)}$, $X_j^{(2)}$, $Y_j^{(1)}$, $Y_j^{(2)}$ are standardized otherwise one can switch to $\widetilde{X}_j^{(1)}$, $\widetilde{X}_j^{(2)}$, $\widetilde{Y}_j^{(1)}$, $\widetilde{Y}_j^{(2)}$ defined by (11) without affecting the values of the empirical correlations coefficients R_{n_1} and R_{n_2}.

Proof of (8)

Note that thanks to (2) and the central limit theorem it holds

$$\sqrt{n_1}\, \overline{X}_{n_1}^{(j)} = O_P(1), \quad \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} \Big[(X_i^{(j)})^2 - 1 \Big] = O_P(1), \quad j = 1, 2. \tag{A1}$$

Furthermore, a Taylor expansion of $x \mapsto \sqrt{x}$ at the point 1 and (A1) yield

$$\Big(\frac{1}{n_1} \sum_{i=1}^{n_1} (X_i^{(j)} - \overline{X}_{n_1}^{(1)})^2 \Big)^{1/2} = 1 + \frac{1}{2n_1} \sum_{i=1}^{n_1} \Big[(X_i^{(j)})^2 - 1 \Big] + o_P(\tfrac{1}{\sqrt{n_1}}), \tag{A2}$$

which implies that

$$
\begin{aligned}
D_{n1} &= \left\{1+\frac{1}{2n_1}\sum_{i=1}^{n_1}\Big[(X_i^{(1)})^2-1\Big]\right\}\left\{1+\frac{1}{2n_1}\sum_{i=1}^{n_1}\Big[(X_i^{(2)})^2-1\Big]\right\} \\
&\quad +o_P(\tfrac{1}{\sqrt{n_1}}) \\
&= \frac{1}{2n_1}\sum_{i=1}^{n_1}(X_i^{(1)})^2+\frac{1}{2n_1}\sum_{i=1}^{n_1}(X_i^{(2)})^2+o_P(\tfrac{1}{\sqrt{n_1}}). \qquad \text{(A3)}
\end{aligned}
$$

With the help of (9), (A1) and (A3) one can derive

$$
\begin{aligned}
\sqrt{n_1}(R_{n_1}-\rho_1) &= \frac{1}{D_{n_1}}\sqrt{n_1}\,(E_{n_1}-\rho_1\,D_{n_1}) \\
&= \frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\Big(X_i^{(1)}X_i^{(2)}-\tfrac{\rho_1}{2}(X_i^{(1)})^2-\tfrac{\rho_1}{2}(X_i^{(2)})^2\Big)+o_P(1) \\
&= \frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}Z_i^X+o_P(1). \qquad \text{(A4)}
\end{aligned}
$$

Analogously one can show that

$$
\sqrt{n_2}(R_{n_2}-\rho_2)=\frac{1}{\sqrt{n_2}}\sum_{i=1}^{n_2}Z_i^Y+o_P(1),
$$

which together with (A4) implies (8).

Auxiliary result - Conditional central limit theorem

Let $\{c_{n,i}, i=1,\dots,n\}$ be a triangular array of constants such that

$$
\sum_{i=1}^{n}c_{n,i}^2=1, \quad \sqrt{n}\max_{1\le i\le n}(c_{n,i}-\bar{c}_n)=O(1), \qquad \text{(A5)}
$$

where $\bar{c}_n=\frac{1}{n}\sum_{i=1}^{n}c_{n,i}$. Next, let $\mathbf{X}_n=\{X_{n,i}, i=1,\dots,n\}$ be a triangular array of random variables defined on $(\Omega,\mathcal{A},P)$. Further, recall that $\tau: (\widetilde{\Omega},\widetilde{\mathcal{A}},\widetilde{P})\to\mathcal{S}_n$ is 'a random permutation', that is a random variable that is uniformly distributed on the symmetric group $\mathcal{S}_n$ and independent of the triangular array $\mathbf{X}_n$. Finally, let the symbol d stand for a distance that metrizes weak convergence, e.g. the Levy distance, see p.398 in Dudley (2002).

The following lemma states the conditional (given $\mathbf{X}_n$) central limit theorem for the quantity $\sum_{i=1}^{n}c_{n,\tau(i)}\,X_{n,i}$ in a way that will be useful in the proofs of Theorem 1 and 2. Here ' $\xrightarrow{P}$ ' stands for convergence in probability with respect to the probability measure P.

Lemma 1. *Suppose that* (A5) *holds and the triangular array* $\mathbf{X}_n$ *satisfies*

$$\max_{1\le i\le n} |X_{n,i} - \overline{X}_n| \xrightarrow{P} 0, \tag{A6}$$

$$\sum_{i=1}^{n} (X_{n,i} - \overline{X}_n)^2 \xrightarrow{P} \sigma^2 \in (0,\infty) \tag{A7}$$

as $n \to \infty$*. Then*

$$d\Big(\mathcal{L}\Big(\sum_{i=1}^{n} c_{n,\tau(i)}(X_{n,i} - \overline{X}_n)|\mathbf{X}_n\Big), N(0,\sigma^2)\Big) \xrightarrow{P} 0, \text{ as } n \to \infty. \tag{A8}$$

Proof. Let us first assume that (A6) and (A7) hold almost surely (a.s.), i.e. for all fixed $\omega \in M^c$ with $P(M) = 0$. For a fixed $\omega \in M^c$ define

$$b_n(i) := c_{n,i} \quad \text{and} \quad a_n(i) := \frac{X_{n,i}(\omega) - \overline{X}_n(\omega)}{\sqrt{\sum_{j=1}^{n}(X_{n,j}(\omega) - \overline{X}_n(\omega))^2}}$$

and denote

$$\tau^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (a_n(i) - \bar{a}_n)^2 \sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^2,$$

where $\bar{a}_n = \frac{1}{n}\sum_{i=1}^{n} a_n(i)$ and analogously for $\bar{b}_n$.

Theorem 4 of Hoeffding (1951) states that the quantity

$$S_n := \sum_{i=1}^{n} a_n(i) b_n(\tau(i))$$

is asymptotically normal with zero mean and variance τ^2, provided

$$\lim_{n\to\infty} \frac{\max_{1\le i\le n}(a_n(i) - \bar{a}_n)^2}{\sum_{i=1}^{n}(a_n(i) - \bar{a}_n)^2} \frac{\max_{1\le i\le n}(b_n(i) - \bar{b}_n)^2}{\sum_{i=1}^{n}(b_n(i) - \bar{b}_n)^2} = 0. \tag{A9}$$

As the condition (A9) is satisfied thanks to (A5), (A6) and (A7), one can conclude that

$$\frac{\sum_{i=1}^{n} c_{n,\tau(i)}(X_{n,i}(\omega) - \overline{X}_n(\omega))}{\sum_{i=1}^{n}(X_{n,i}(\omega) - \overline{X}_n(\omega))^2} \tag{A10}$$

converges to a standard normal distribution $N(0,1)$. As for all fixed $\omega \in M^c$ one has $\sum_{i=1}^n (X_{n,i}(\omega) - \overline{X}_n(\omega))^2 \to \sigma^2$, the convergence (A8) holds even a.s. (with respect to P). Now the desired result follows from the subsequence principle for convergence in probability, see e.g. Theorem 9.2.1. in Dudley (2002).

$\square$

In the sequel we will apply Lemma 1 with the following two triangular arrays of coefficients:

$$1. \qquad c_{n,i} := \sqrt{\frac{n_1 n_2}{n}} \cdot \begin{cases} \frac{1}{n_1}, & \text{for } 1 \le i \le n_1, \\ -\frac{1}{n_2}, & \text{for } n_1 < i \le n. \end{cases} \tag{A11}$$

$$2. \qquad d_{n,i} := \sqrt{\frac{n_1 n}{n_2}} \cdot \begin{cases} \frac{1}{n_1}, & \text{for } 1 \le i \le n_1, \\ 0, & \text{for } n_1 < i \le n. \end{cases} \tag{A12}$$

Proof of Theorem 1

Note that the statistic $T_n = a_n(R_{n2} - R_{n1})$ can be written as a function of the standardized observations $\mathbb{S}_n = (S_{n,1}, \ldots, S_{n,n})$. For a random permutation τ put $\mathbb{S}_{n,\tau} = (S_{n,\tau(1)}, \ldots, S_{n,\tau(n)})$.

Assume for a moment that under H_0 it holds that

$$d\big(\mathcal{L}(T_n(\mathbb{S}_{n,\tau})|\mathbb{S}_n), N(0, p\,\sigma_X^2 + q\,\sigma_Y^2)\big) \xrightarrow{P} 0, \quad \text{as } n \to \infty, \tag{A13}$$

where σ_1^X and σ_1^Y are defined below the equation (12). As by the assumptions of Theorem 1 one gets $p\,\sigma_X^2 + q\,\sigma_Y^2 = q\,\sigma_X^2 + p\,\sigma_Y^2$, Lemma 1 of Janssen and Pauls (2003) and the asymptotic normality of T_n given by (12) imply the statements of Theorem 1.

Suppose for a moment that the following analogy of (8) holds

$$T_n(\mathbb{S}_{n,\tau}) = \sum_{i=1}^n c_{n,i} \widetilde{Z}_{n,\tau(i)} + o_{P \otimes \widetilde{P}}(1), \tag{A14}$$

where $\{\widetilde{Z}_{n,i}\}$ were defined in (15) and $\{c_{n,i}\}$ in (A11). Since the non-vanishing linear part of (A14) is equal in distribution to $\sum_{i=1}^n c_{n,\tau(i)} \widetilde{Z}_{n,i}$ (conditionally on $\mathbb{S}_n$) it is straightforward to verify (A13) using the expansion (A14) and Lemma 1. Thus it remains to prove (A14).

When verifying (A14) we proceed as in the proof of (8). Thus we need to show that for $j = 1, 2$

$$\frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} S_{n,\tau(i)}^{(j)} = O_{P \otimes \widetilde{P}}(1) \tag{A15}$$

and

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1}\Big[(S_{n,\tau(i)}^{(j)})^2-1\Big]=O_{P\otimes\widetilde{P}}(1). \tag{A16}$$

Note that as $\sum_{i=1}^{n} S_{n,i}^{(j)}=0$ one has the following equality of distributions

$$\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1} S_{n,\tau(i)}^{(j)} \overset{d}{=} \frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1} d_{n,\tau(i)}(X_{n,i}-\bar{X}_n),$$

where the coefficients $\{d_{n,i}\}$ are defined in (A12) and $X_{n,i}=\sqrt{\frac{n_2}{n_1 n}}\,S_{n,i}^{(j)}$. Condition (A6) is satisfied thanks to the law of large numbers and assumption (2). Moreover, we have

$$\begin{aligned}
\sum_{i=1}^{n}(X_{n,i}-\overline{X}_n)^2 &= \frac{n_2}{n_1 n}\sum_{i=1}^{n}(S_{n,i}^{(j)})^2 \\
&= \frac{n_2}{n_1 n}\left(\sum_{i=1}^{n_1}(S_{n,i}^{(j)})^2+\sum_{i=n_1+1}^{n}(S_{n,i}^{(j)})^2\right) \\
&= \frac{n_2}{n}+\frac{n_2^2}{n_1\,n}=\frac{n_2}{n_1}\to\frac{1-p}{p},
\end{aligned}$$

thus (A7) also holds. Now Lemma 1 yields that the quantity $\frac{1}{\sqrt{n_1}}\sum_{i=1}^{n_1} S_{n,\tau(i)}^{(j)}$ is asymptotically normal in the sense of (A8), which further implies (A15).

Analogously one can prove (A16) by taking $X_{n,i}=\sqrt{\frac{n_2}{n_1 n}}\,(S_{n,i}^{(j)})^2$ and utilizing that $\frac{1}{n}\sum_{i=1}^{n}(S_{n,i}^{(j)})^2=1$.

Finally, by the same reasoning as in the proof of (8) one can conclude the proof of (A14) and thus the proof of the theorem is finished. □

Proof of Theorems 2(a)

First note that by the law of large numbers

$$V_n^2 \xrightarrow{P} q\,\sigma_X^2+p\,\sigma_Y^2 \quad \text{as } n\to\infty, \tag{A17}$$

where V_n^2 is the variance estimator for the test statistic T_n defined by (14).

Suppose for a moment that

$$V_n^2(\mathbb{S}_{n,\tau}^{*}) \xrightarrow{P\otimes\widetilde{P}} q\,\sigma_X^2+p\,\sigma_Y^2 \quad \text{as } n\to\infty, \tag{A18}$$

where $\mathbb{S}^*_{n,\tau} = (S^*_{n,\tau(1)}, \dots, S^*_{n,\tau(n)})$. Then Slutsky's Lemma and the already proved result (A13) for the statistic T_n imply that

$$d\big(\mathcal{L}(\widetilde{T}_n(\mathbb{S}^*_{n,\tau})|\mathbb{S}_n), N(0,1)\big) \xrightarrow{P} 0 \quad \text{as } n \to \infty. \tag{A19}$$

Now Lemma 1 of Janssen and Pauls (2003) together with (A19) concludes the statement (a) of the theorem.

Thus it remains to verify (A18). After some algebraic manipulations this problem reduces to proving that

$$\frac{1}{n_1} \sum_{i=1}^{n_1} (S^{(1)}_{n,\tau(i)})^{l_1} (S^{(2)}_{n,\tau(i)})^{l_2} = p\, \xi_X^{(l_1,l_2)} + q\, \xi_Y^{(l_1,l_2)} + o_{P\otimes\widetilde{P}}(1), \tag{A20}$$

where

$$\xi_X^{(j,k)} := \mathrm{E}\left[(X_1^{(1)})^j (X_1^{(2)})^k\right], \qquad \xi_Y^{(j,k)} := \mathrm{E}\left[(Y_1^{(1)})^j (Y_1^{(2)})^k\right].$$

and $(l_1, l_2) \in M = \{(i,j) : i,j \in \{0,1,2,3,4\}, i+j \le 4\}$.

To simplify the notation for fixed $(l_1,l_2) \in M$ put $W_{n,i} = (S^{(1)}_{n,i})^{l_1} (S^{(2)}_{n,i})^{l_2}$. Thus our aim is to prove that

$$\frac{1}{n_1} \sum_{i=1}^{n_1} W_{n,\tau(i)} = p\, \xi_X^{(l_1,l_2)} + q\, \xi_Y^{(l_1,l_2)} + o_{P\otimes\widetilde{P}}(1). \tag{A21}$$

To prove (A21) we will proceed similarly as in the proof of Lemma 4.1 of Janssen (1997). For $C > 0$ introduce the truncated observations

$$W^C_{n,i} = \begin{cases} W_{n,i}\, \mathbb{I}\{|X_i^{(1)}| \le C\,,\, |X_i^{(2)}| \le C\}, & i = 1, \dots, n_1, \\ W_{n,n_1+i}\, \mathbb{I}\{|Y_i^{(1)}| \le C\,,\, |Y_i^{(2)}| \le C\}, & i = 1, \dots, n_2. \end{cases}$$

Note that thanks to

$$\overline{X}^{(j)}_{n_1} \xrightarrow{P} 0, \quad \frac{1}{n_1} \sum_{i=1}^{n_1} (X_i^{(j)})^2 \xrightarrow{P} 1, \quad \overline{Y}^{(j)}_{n_2} \xrightarrow{P} 0, \quad \frac{1}{n_2} \sum_{i=1}^{n_2} (Y_i^{(j)})^2 \xrightarrow{P} 1,$$

one can assume that $|W^{C(j)}_{n,i}| \le (4\,C)^{l_1+l_2}$ holds on a set of asymptotic probability 1. Further define

$$\xi_{X^C}^{(l_1,l_2)} = \mathrm{E}\left[(X_1^{C(1)})^{l_1}\, (X_1^{C(2)})^{l_2}\right], \quad \text{where } X_1^{C(j)} = X_1^{(j)}\, \mathbb{I}\{|X_1^{(j)}| \le C\},$$

for $j = 1, 2$. Analogously define $\xi_{Y^C}^{(l_1,l_2)}$.

Now everything is ready to decompose

$$
\begin{aligned}
&\frac{1}{n_1}\sum_{i=1}^{n_1} W_{n,\tau(i)} - p\xi_X^{(l_1,l_2)} - q\xi_Y^{(l_1,l_2)} \\
&\quad = \frac{1}{n_1}\sum_{i=1}^{n_1} W_{n,\tau(i)} - \frac{1}{n_1}\sum_{i=1}^{n_1} W^C_{n,\tau(i)} && \text{(A22)}\\
&\qquad + \frac{1}{n_1}\sum_{i=1}^{n_1} W^C_{n,\tau(i)} - \frac{1}{n}\sum_{i=1}^{n} W^C_{n,i} && \text{(A23)}\\
&\qquad + \frac{1}{n}\sum_{i=1}^{n} W^C_{n,i} - p\xi_{X^C}^{(l_1,l_2)} - q\xi_{Y^C}^{(l_1,l_2)} && \text{(A24)}\\
&\qquad + p\,\xi_{X^C}^{(l_1,l_2)} + q\,\xi_{Y^C}^{(l_1,l_2)} - p\,\xi_X^{(l_1,l_2)} - q\,\xi_Y^{(l_1,l_2)}. && \text{(A25)}
\end{aligned}
$$

Thanks to assumption (2) the fourth term (A25) can be made arbitrarily small by taking C large enough. Next, the third term (A24) is $o_P(1)$ by applying the law of large numbers. Further, as one can suppose that the random variables $\{W^C_{n,i}, i \leq n\}$ are bounded (see the note below the definition of $W^C_{n,i}$), the second term (A23) is $o_{\widetilde{P}}(1)$ by Chebyshev's inequality. Finally, with the help of the Markov's inequality, and the law of large numbers the first term (A22) can be handled as

$$
\begin{aligned}
&\widetilde{\mathrm{P}}\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} W_{n,\tau(i)} - \frac{1}{n_1}\sum_{i=1}^{n_1} W^C_{n,\tau(i)}\right| > \varepsilon\right] \\
&\quad \leq \frac{1}{\varepsilon}\,\mathrm{E}_{\widetilde{P}}\left|\frac{1}{n_1}\sum_{i=1}^{n_1} W_{n,\tau(i)} - \frac{1}{n_1}\sum_{i=1}^{n_1} W^C_{n,\tau(i)}\right| \leq \frac{1}{n\,\varepsilon}\sum_{i=1}^{n}\left|W_{n,i} - W^C_{n,i}\right| \\
&\quad \leq \frac{1}{n\,\varepsilon}\left(\sum_{i=1}^{n_1}\left|(X_i^{(1)})^{l_1}(X_i^{(2)})^{l_2} - (X_i^{C(1)})^{l_1}(X_i^{C(2)})^{l_2}\right|\right.\\
&\qquad\qquad \left. + \sum_{i=1}^{n_2}\left|(Y_i^{(1)})^{l_1}(Y_i^{(2)})^{l_2} - (Y_i^{C(1)})^{l_1}(Y_i^{C(2)})^{l_2}\right|\right) + o_P(1) \\
&\quad \leq p\,\mathrm{E}\left|(X_1^{(1)})^{l_1}(X_1^{(2)})^{l_2} - (X_1^{C(1)})^{l_1}(X_1^{C(2)})^{l_2}\right| \\
&\qquad\qquad + q\,\mathrm{E}\left|(Y_1^{(1)})^{l_1}(Y_1^{(2)})^{l_2} - (Y_1^{C(1)})^{l_1}(Y_1^{C(2)})^{l_2}\right| + o_P(1),
\end{aligned}
$$

which can be made arbitrarily small by taking C large enough. Thus we have proved (A21), which concludes the proof of part (a) of Theorem 2. □

Proof of Theorems 2 (b)

Note that as $\rho_1 \neq \rho_2$ the limits of the variances V_n^2 and $V_n^2(\mathbb{S}^*_{n,\tau})$ are generally different. For the variance estimator V_n^2 we still have the convergence (A17). Next, by the same arguments as in the proof of Theorems 2(a) one can show that

$$V_n^2(\mathbb{S}_{n,\tau}) = p\,\mathrm{var}\,(\widetilde{Z}^X) + q\,\mathrm{var}\,(\widetilde{Z}^Y) + o_{P\otimes\widetilde{P}}(1), \tag{A26}$$

where

$$\begin{aligned}
\widetilde{Z}^X &= \widetilde{X}_1^{(1)}\widetilde{X}_1^{(2)} - \tfrac{\rho}{2}\big((\widetilde{X}_1^{(1)})^2 + (\widetilde{X}_1^{(2)})^2\big),\\
\widetilde{Z}^Y &= \widetilde{Y}_1^{(1)}\widetilde{Y}_1^{(2)} - \tfrac{\rho}{2}\big((\widetilde{Y}_1^{(1)})^2 + (\widetilde{Y}_1^{(2)})^2\big),
\end{aligned}$$

with $\rho := p\rho_1 + q\rho_2$.

Note that all the above variances are positive. This can be justified by the fact that for $|a| \leq 1$ the equation

$$xy - \tfrac{a}{2}(x^2 + y^2) = 0,$$

has the only solutions $x = cy$. Thus for any two non-degenerate zero mean random variables U and V such that $|\,\mathrm{corr}(U, V)| < 1$

$$\mathrm{var}\,\{U\,V - \tfrac{a}{2}(U^2 + V^2)\} > 0 \quad \text{for } |a| \leq 1.$$

Finally (A17), (A26), together with $\frac{1}{a_n}T_n \xrightarrow{P} \rho_1 - \rho_2$, and $\frac{1}{a_n}T_n(\mathbb{S}_{n,\tau}) = o_{P\otimes\widetilde{P}}(1)$ imply the consistency of the test $\widetilde{\varphi}^*_n$, which completes the proof. $\square$

Acknowledgements

The authors would like to thank Lutz Duembgen for helpful comments and remarks. The authors are also grateful to two reviewers for their very helpful comments which led to a considerable improvement of the original version of the paper. The work of the co-Editor is also very appreciated.

The work of the first author was supported by the grants MSM 002162839 amd GACR P201/11/P290. The second author was supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD).

References

Argaç, D., 2003. Testing for homogeneity in a general one-way classification with fixed effects: Power simulations and comparative study. Comput. Stat. Data Anal. 44 (4), 603–612.

Brombin, C., Salmaso, L., Ferronato, G., Galzignato, P.-F., 2011. Multi-aspect procedures for paired data with application to biometric morphing. Communications in Statistics - Simulation and Computation 40 (1), 1–12.

Davison, A. C., Hinkley, D. V., 1997. Bootstrap Methods and their Application. Cambridge University Press, New York.

Dudley, R., 2002. Real analysis and probability. Cambridge Studies in Advanced Mathematics. Cambridge University Press.

Edgington, E. S., Onghena, P., 2007. Randomization tests. Chapman & Hall.

Good, P. I., 2005. Permutation, parametric and bootstrap tests of hypotheses. 3rd ed. Springer Series in Statistics, New York.

Hoeffding, W., 1951. A combinatorial central limit theorem. Ann. Math. Stat. 22, 558–566.

Janssen, A., 1997. Studentized permutation tests for non-i.i.d hypotheses and the generalized Behrens-Fisher problem. Statist. Probab. Lett. 36, 9–21.

Janssen, A., Pauls, T., 2003. How do bootstrap and permutation tests work? Ann. Statist. 31 (3), 768–806.

Jennrich, R., 1970. An asymptotic χ^2 test for the equality of two correlation matrices. Journal of the American Statistical Association 65 (2), 904–912.

Kullback, S., 1967. On testing correlation matrices. Applied Statistics, 80–85.

Lehmann, E., Romano, J., 2005. Testing statistical hypotheses. Springer Verlag.

Neuhaus, G., 1993. Conditional rank tests for the two-sample problem under random censorship. Ann. Statist. 21 (4), 1760–1779.

Paul, S., 1989. Test for the equality of several correlation coefficients. Can. J. Stat. 17 (2), 217–227.

Pearson, E., Wilks, S., 1933. Methods of statistical analysis appropriate for k samples of two variables. Biometrika 25, 353–378.

Pesarin, F., Salmaso, L., 2010. Permutation Tests for Complex Data. Wiley Series in Probability and Statistics.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Romano, J., 1990. On the behavior of randomization tests without a group invariance assumption. Journal of the American Statistical Association 85 (411), 686–692.

Sakaori, F., 2002. Permutation test for equality of correlation coefficients in two populations. Communications in Statistics-Simulation and Computation 31 (4), 641–651.

Van der Vaart, A., 2000. Asymptotic statistics. Cambridge Univ. Press.