STATISTICS IN TRANSITION new series, December 2019
Vol. 20, No. 4, pp. 153–166, DOI 10.21307/stattrans-2019-039
Submitted – 04.03.2019; Paper ready for publication – 27.09.2019

ON PERMUTATION LOCATION–SCALE TESTS

Dominika Polko-Zając [1]

ABSTRACT

Statisticians are constantly looking for methods of statistical inference that would be both effective and would require meeting as few assumptions as possible. Permutation tests seem to fit here, as using them makes it possible to perform statistical inference in situations where classical parametric tests do not work. Permutation tests appear to be comparably powerful to parametric tests, but require meeting fewer assumptions, e.g. regarding the size of the sample or the from of distribution of the tested variable in a population. The presented tests make it possible to verify the overall hypothesis about the identity of both location and scale parameters in the studied populations. In literature, the Lepage test and the Cucconi test are most often referred to in this context. The paper considers various forms of test statistics, and presents a simulation study carried out to determine the size and power of the tests under normality. As the study demonstrated, the advantage of the proposed method is that it can be applied to small-size samples. A nonparametric, complex procedure was used to assess the overall ASL (achieved significance level) value by applying the permutation principle. For comparative purposes, the results for the permutation Lepage test and the permutation Cucconi test are also presented.

Key words: permutation tests, comparing populations, test power, the Lepage test, the Cucconi test.

1. Introduction

Comparing populations most frequently refers to a comparison of the characteristics of these populations. If it is assumed that population distributions differ only in the central tendency, there are various parametric and nonparametric tests to verify this hypothesis. Many authors undertake to study the power and size of tests for the significance of differences between means or medians in two or more populations, using for this purpose simulation methods based on bootstrap or permutation tests (Janssen and Pauls, 2005; Chang and Pal, 2008; Kończak, 2016; Anderson et al., 2017). The problem of comparing variances in populations is also common in research. For example, comparative studies using simulations were conducted by Hall (1972), Geng, Wang and Miller (1979), Keselman, Games and Clinch (1979), Conover, Johnson and Johnson

[1] Department of Statistics, Econometrics and Mathematics, University of Economics in Katowice, Poland. E-mail: dominika.polko@ue.katowice.pl. ORCID ID: https://orcid.org/0000-0003-4098-6647.

(1981), Balakrishnan and Ma (1990), Lim and Loh (1996), Marozzi (2011) and Gogoi and Gogoi (2017).

Pesarin (2001) initiated the approach for the nonparametric testing problem. He considered reducing the scope of the null hypothesis by splitting it into several partial hypotheses. This nonparametric approach is to perform some reasonable tests for each individual partial hypothesis and combine their results with a chosen combining function. A multi–aspect test to location problem was considered in works by Marozzi (2004), Marozzi (2007) or Salmaso and Solari (2005). Nonparametric combination procedure to asses overall *ASL* (*achieved significance level*) value is very useful in the scale problem too (Marozzi 2012a, 2012b).

It is more complicated to test differences between both location parameters and scale parameters of the distribution in the populations studied. A need of simultaneously detecting location and scale changes arises in many areas, for example in financial matters in stock prices analysis (Lunde and Timmermann, 2004), in the analysis of production processes, for example, testing of the process stability (Park, 2015a), climate dynamics analyses (Yonetani and Gordon, 2001) or biomedical researches (Muccioli, et al., 1996).

Lepage (1971) initiated this topic with his proposal by combining the Wilcoxon rank sum and Ansari–Bradley's test statistics for location and scale parameters. A test based on Lepage's proposal but using Mood's test statistic for the scale parameter was presented by Duran et al. (1976). Later, Lepage's procedure was reviewed and discussed extensively by many authors (Murakami, 2007; Neuhauser, Leuchs and Ball, 2011). Marozzi (2008) considered the problem of location and scale using a nonparametric combination procedure proposed by Pesarin (2001). All the reviewed and compared by a simulation study test statistics used quadratic forms and allow one to consider only two–sided alternatives. Park (2015b) excluded the use of the quadratic form for the test statistics to accommodate various types of alternatives. The proposition described in this article also enables the formulation of any types of alternative hypotheses. The purpose of this research is to present several statistical test proposals for joint comparison of location and scale parameters in two populations using a permutation procedure for a multi–aspect testing approach.

The rest of the paper is organized as follows. In Section 2 the research problem is formally defined and two tests known in the literature for simultaneous testing location and scale parameters are presented. In Section 3 the nonparametric combination procedure for location–scale problem is characterized. In Section 4 several test statistics for a joint comparison of the location and scale parameters in two populations using a nonparametric, permutation procedure to assess *ASL* values are proposed. This Section also contains a simulation comparison of their size and power under normality. There are two cases considered in simulations: both partial alternative hypotheses are one– or two–sided. In Section 5 concluding remarks are presented.

2. Simultaneous tests for the location–scale problem

In order to discuss the location–scale problem, let observations $x_{11},...,x_{1n_1}$ and $x_{21},...,x_{2n_2}$ be random samples taken from populations with distribution functions F_1 and F_2 respectively. Populations are of continuous distributions F_i for $i = 1$, 2 with unknown parameters. The null hypothesis of comparing two populations is in the form of $H_0 : F_1(x) = F_2(x)$. In the paper, the location–scale problem is considered where μ_1, μ_2 and σ_1, σ_2 are locations and scale parameters of populations 1 and 2 respectively. According to this notation, the null hypothesis can be also written as

$$H_0 : \mu_1 = \mu_2 \wedge \sigma_1 = \sigma_2, \tag{1}$$

versus alternative hypothesis

$$H_1 : \mu_1 \neq \mu_2 \vee \sigma_1 \neq \sigma_2. \tag{2}$$

In the literature, authors most often refer to the Lepage test. However, you can find another test to verify the same hypothesis, proposed earlier, but not so well known Cucconi test (Bonnini, et al., 2014). The Cucconi test (Cucconi, 1968) used in the situations of finding differences in the location and scale parameters uses the statistic of the form (Marozzi, 2009)

$$C = \frac{U^2 + V^2 - 2\rho UV}{2(1-\rho^2)}, \tag{3}$$

where

$$U = \frac{6\sum_{i=1}^{n_1} R_{1i}^2 - n_1(n+1)(2n+1)}{\sqrt{n_1 n_2 (n+1)(2n+1)(8n+11)/5}},$$

$$V = \frac{6\sum_{i=1}^{n_1} (n+1-R_{1i})^2 - n_1(n+1)(2n+1)}{\sqrt{n_1 n_2 (n+1)(2n+1)(8n+11)/5}},$$

$n = n_1 + n_2$,

R_{ji} – rank of x_{ji} in pooled sample $x = (x_1, x_2)$,

and $\rho = \frac{2(n^2-4)}{(2n+1)(8n+11)} - 1$.

Hypothesis H_0 is rejected if C>-lnα, where α is the test size (Marozzi, 2009).

The second test, the Lepage test (1971), refers to the merger of two test statistics. This test is a combination of the Wilcoxon–Mann–Whitney (Mann and Whitney, 1947; Wilcoxon 1949) and Ansari–Bradley (Ansari and Bradley, 1960) test statistics

$$L = \frac{(W - E_0(W))^2}{V_0(W)} + \frac{(A - E_0(A))^2}{V_0(A)} = \widetilde{W}^2 + \widetilde{A}^2, \tag{4}$$

where

W – Wilcoxon–Mann–Whitney test statistics,

A – Ansari–Bradley test statistics,

$E_0(W) = n_1(n+1)/2$, $V_0(W) = n_1 n_2 (n+1)/12$,

when n is even: $E_0(A) = n_1(n+2)/4$, $V_0(A) = n_1 n_2 (n+2)(n-2)/48/(n-1)$,

when n is odd: $E_0(A) = n_1(n+1)^2/4/n$, $V_0(A) = n_1 n_2 (n+1)(n^2+3)/48/n^2$,

$\widetilde{W}$ – Wilcoxon–Mann–Whitney standardized test statistics,

$\widetilde{A}$ – Ansari–Bradley standardized test statistics.

Hypothesis H_0 is rejected if the calculated value of the test statistic exceeds the critical value of the test. Tables for the Lepage test can be found in Lepage (1973).

3. Nonparametric combination procedures

The problem of testing complex hypotheses can also be considered as proposed by Pesarin (2001). When the test concerns the location–scale testing problem then two partial hypotheses are taken into account. The null hypothesis in the form of (1) can be written differently as

$$H_0 : H_0^{(1)} \wedge H_0^{(2)} \tag{5}$$

and the corresponding decomposition is

$$H_0^{(1)} : \mu_1 = \mu_2 \text{ and } H_0^{(2)} : \sigma_1 = \sigma_2 . \tag{6}$$

An alternative hypothesis, which is a negation of the null hypothesis, can then be written as

$$H_1 : H_1^{(1)} \vee H_1^{(2)}, \tag{7}$$

where

$$H_1^{(1)} : \mu_1 \neq \mu_2 , \quad H_1^{(2)} : \sigma_1 \neq \sigma_2 . \tag{8}$$

The paper considers a simulation approach based on the permutations of a data set. A nonparametric, complex procedure was used to assess the overall *ASL* (*achieved significance level*) value. The procedure for testing the null hypothesis versus the alternative hypothesis consists of two steps. First, each of the partial null hypotheses is tested. Then, the results of the first step are jointly managed to solve the general problem (Marozzi, 2008).

In the first stage of separate testing of each of the considered partial null hypotheses, *ASL* values are estimated following the traditional permutation method used during the verification of a single parameter hypothesis, i.e.:

1. Assume the level of significance α.
2. Calculate the value of statistic for the sample data (T_0).
3. Perform the permutations of variable *N*–times and calculate the statistic test value (T_k) for each permutation.
4. On the basis of the empirical distribution of statistic, the *ASL* value is determined.

Regarding location–scale testing, two partial aspects may be emphasized. Two permutation tests are performed and an estimate of two *ASL* values are obtained: the first for the equality test of mean or median parameters, the second for the equality test of scale parameters of the form

$$A\hat{S}L_{T^{(1)}}\left(T_0^{(1)}\right) = \frac{0{,}5 + \sum_{k=1}^{N} I\left(\left|T_k^{(1)}\right| \geq \left|T_0^{(1)}\right|\right)}{N+1} \tag{9}$$

and

$$A\hat{S}L_{T^{(2)}}\left(T_0^{(2)}\right) = \frac{0{,}5 + \sum_{k=1}^{N} I\left(\left|T_k^{(2)}\right| \geq \left|T_0^{(2)}\right|\right)}{N+1}. \tag{10}$$

where *I*(.) denotes the indicator function.

With respect to standard permutation *ASL* estimation, 0.5 and 1 are added to the numerator and denominator of the fraction, respectively. The reason is to obtain estimated *ASL* values in open interval $(0{,}1)$ avoiding computational problems, which may arise in the second step of the nonparametric procedure. However, since large *N* is used, this correction is practically irrelevant (Marozzi, 2008).

The second step of the nonparametric procedure of statistical inference includes calculation of the overall *ASL* value using the combining function (Pesarin, 2001)

$$_{\varphi 12}T = \varphi\left(ASL_{T^{(1)}}, ASL_{T^{(2)}}\right).$$

There are many forms of combining functions for determining the overall *ASL* value, although the authors the most often used combining functions:

- the Fisher omnibus combining function (Fisher, 1932)
  $C^{(F)} = -2\left[\log\left(A\hat{S}L_{T^{(1)}}\right) + \log\left(A\hat{S}L_{T^{(2)}}\right)\right],$

- the Liptak combining function (Liptak 1958) $C^{(L)} = \Phi^{-1}\left(1 - A\hat{S}L_{T^{(1)}}\right) + \Phi^{-1}\left(1 - A\hat{S}L_{T^{(2)}}\right)$, where Φ denotes the standard normal distribution function,
- the Tippet combining function (Tippet, 1931) $C^{(T)} = \max\left\{1 - A\hat{S}L_{T^{(1)}}, 1 - A\hat{S}L_{T^{(2)}}\right\}$.

The observed statistics value for the sample data can be determined as

$$_{\varphi 12}T_0 = \varphi\left(A\hat{S}L_{T^{(1)}}\left(T_0^{(1)}\right), A\hat{S}L_{T^{(2)}}\left(T_0^{(2)}\right)\right), \tag{11}$$

and its distribution is determined on the basis of the same permutations of the first step of this procedure, for example the *k*–th permutation value of statistics is computed

$$_{\varphi 12}T_k = \varphi\left(A\hat{S}L_{T^{(1)}}\left(T_k^{(1)}\right), A\hat{S}L_{T^{(2)}}\left(T_k^{(2)}\right)\right). \tag{12}$$

Overall *ASL* value of the test is estimated by using the formula

$$A\hat{S}L_{_{\varphi 12}T} = \frac{\sum_{k=1}^{N} I\left(_{\varphi 12}T_k \geq {}_{\varphi 12}T_0\right)}{N}. \tag{13}$$

where *I*(.) denotes the indicator function.

4. Monte Carlo study

Most often, the statistical inference concerns situations where there are differences between the considered populations without indicating the nature of this difference. An alternative hypothesis of form (2) is then considered. Thanks to permutation tests, it is also possible to consider one–sided alternative hypotheses, for example:

$$H_1^{(1)}: \mu_1 > \mu_2 \text{ or } H_1^{(2)}: \sigma_1 > \sigma_2.$$

The study considered various forms of test statistics (Table 1). The simulations consisted of calculating the size and power of the presented tests using a complex, nonparametric method of testing the location and scale parameters. All 1–8 models were used in the simulation study when partial alternative two–sided hypotheses were considered. To verify the null hypothesis, when partial alternative hypotheses were one–sided hypotheses, models 1–5 were used. Model 6 considers the form of test statistics included in the combination of statistics used in the Lepage test. For comparative purposes, the results for the permutation Lepage test (model 7) and permutation Cucconi test (model 8) were also included when alternative two–sided hypotheses were considered. The nonparametric combination procedure for the estimated overall *ASL* value was used when considering models 1–6.

In the simulation study samples taken from normal distribution with $n_1 = 10, n_2 = 15$ sample sizes were considered. Three situations were analysed:

a) $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 = 1$,

b) $\mu_1 - \mu_2 = 0$ and $\sigma_1 / \sigma_2 > 1$,

c) $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 > 1$.

Parameters of the distribution from which the second sample was taken are $\mu_2 = 0$ and $\sigma_2 = 1$, whereas parameters of the distribution from which the first sample was taken are defined as follows:

a) if $\mu_1 - \mu_2 > 0$ then $\mu_1 \in (0.2, 1.6)$ with the increment 0.2 and $\sigma_1 = 1$,

b) if $\sigma_1 / \sigma_2 > 1$ then $\sigma_1 \in (1.2, 2.6)$ with the increment 0.2 and $\mu_1 = 0$,

c) if $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 > 1$ then parameters of the distribution (μ_1, σ_1) equal from (0.2,1.2) to (1.6,2.6) with the increment 0.2 for each parameter.

Table 1. Statistics used in simulation study

Model	Statistics $T^{(1)}$	Statistics $T^{(2)}$
1	$T_1^{(1)} = \overline{x}_1 - \overline{x}_2$	$T_1^{(2)} = \dfrac{s_1^2}{s_2^2}$
2	$T_2^{(1)} = m_1 - m_2$	$T_2^{(2)} = \dfrac{R_1}{R_2}$,
3	$T_3^{(1)} = W$	$T_3^{(2)} = \dfrac{R_1}{R_2}$,
4	$T_4^{(1)} = W$	$T_4^{(2)} = M$,
5	$T_5^{(1)} = W$	$T_5^{(2)} = OB$,
6	$T_6^{(1)} = \widetilde{W}^2$	$T_6^{(2)} = \widetilde{A}^2$
7	$T_7 = L$	
8	$T_8 = C$	

where:

$\overline{x}_1, \overline{x}_2$ – sample means from first and second population respectively,

m_1, m_2 – sample medians from first and second population respectively,

R_1, R_2 – sample ranges from first and second population respectively,

W – Wilcoxon–Mann–Whitney test statistics,

M – Mood test statistics (Mood, 1954),

OB – O'Brien test statistics (O'Brien, 1979),

$\widetilde{W}$ – Wilcoxon–Mann–Whitney standardized test statistics,

$\widetilde{A}$ – Ansari–Bradley standardized test statistics,
L – Lepage test statistics (4),
C – Cucconi test statistics (3).

Table 2. Size and power estimates when $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 = 1$, $\alpha = 0.05$, for samples $n_1 = 10, n_2 = 15$ (two–sided alternative hypotheses)

Model	Distribution parameters (μ_1, σ_1)								
	$(0,1)$	$(0.2,1)$	$(0.4,1)$	$(0.6,1)$	$(0.8,1)$	$(1,1)$	$(1.2,1)$	$(1.4,1)$	$(1.6,1)$
1	0.046	0.060	0.124	0.240	0.375	0.527	0.684	0.793	0.895
2	0.048	0.065	0.114	0.206	0.307	0.430	0.570	0.689	0.816
3	0.060	0.096	0.188	0.315	0.467	0.626	0.774	0.863	0.940
4	0.054	0.094	0.163	0.295	0.450	0.566	0.734	0.831	0.922
5	0.060	0.104	0.164	0.307	0.464	0.605	0.765	0.849	0.930
6	0.057	0.069	0.111	0.207	0.343	0.502	0.649	0.765	0.889
7	0.052	0.065	0.104	0.210	0.349	0.510	0.653	0.783	0.894
8	0.055	0.072	0.105	0.207	0.347	0.508	0.662	0.769	0.890

Source: Own calculation in R program.

Table 3. Power estimates when $\mu_1 - \mu_2 = 0$ and $\sigma_1 / \sigma_2 > 1$, $\alpha = 0.05$, for samples $n_1 = 10, n_2 = 15$, (two–sided alternative hypotheses)

Model	Distribution parameters (μ_1, σ_1)							
	$(0,1.2)$	$(0,1.4)$	$(0,1.6)$	$(0,1.8)$	$(0,2)$	$(0,2.2)$	$(0,2.4)$	$(0,2.6)$
1	0.115	0.232	0.308	0.440	0.548	0.635	0.750	0.784
2	0.112	0.203	0.305	0.417	0.498	0.604	0.713	0.752
3	0.092	0.176	0.280	0.396	0.455	0.558	0.685	0.722
4	0.079	0.126	0.210	0.282	0.365	0.443	0.527	0.596
5	0.085	0.167	0.276	0.401	0.478	0.580	0.681	0.718
6	0.068	0.133	0.159	0.263	0.316	0.386	0.477	0.545
7	0.077	0.148	0.210	0.307	0.390	0.463	0.573	0.647
8	0.069	0.132	0.172	0.263	0.327	0.399	0.484	0.550

Source: Own calculation in R program.

Table 4. Power estimates when $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 > 1$, $\alpha = 0.05$, for samples $n_1 = 10, n_2 = 15$, (two–sided alternative hypotheses)

Model	Distribution parameters (μ_1, σ_1)							
	$(0.2,1.2)$	$(0.4,1.4)$	$(0.6,1.6)$	$(0.8,1.8)$	$(1,2)$	$(1.2,2.2)$	$(1.4,2.4)$	$(1.6,2.6)$
1	0.145	0.269	0.463	0.624	0.742	0.840	0.874	0.930
2	0.127	0.251	0.422	0.579	0.681	0.792	0.850	0.896
3	0.165	0.309	0.471	0.651	0.742	0.833	0.865	0.920
4	0.124	0.239	0.374	0.516	0.646	0.762	0.812	0.846
5	0.131	0.288	0.469	0.642	0.762	0.859	0.889	0.935
6	0.090	0.161	0.263	0.384	0.508	0.600	0.691	0.739
7	0.106	0.191	0.319	0.424	0.577	0.659	0.748	0.785
8	0.095	0.159	0.295	0.375	0.502	0.604	0.683	0.738

Source: Own calculation in R program.

Table 5. Size and power estimates when $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 = 1$, $\alpha = 0.05$, for samples $n_1 = 10, n_2 = 15$ (one–sided alternative hypotheses)

Model	Distribution parameters (μ_1, σ_1)								
	$(0,1)$	$(0.2,1)$	$(0.4,1)$	$(0.6,1)$	$(0.8,1)$	$(1,1)$	$(1.2,1)$	$(1.4,1)$	$(1.6,1)$
1	0.047	0.105	0.179	0.293	0.495	0.651	0.771	0.885	0.954
2	0.042	0.096	0.166	0.261	0.402	0.568	0.691	0.830	0.925
3	0.050	0.099	0.169	0.295	0.460	0.645	0.763	0.878	0.953
4	0.047	0.100	0.178	0.282	0.454	0.634	0.752	0.875	0.947
5	0.043	0.097	0.162	0.277	0.447	0.627	0.754	0.870	0.948

Source: Own calculation in R program.

For each of 1000 Monte Carlo simulations, 1000 random permutations of variables and the nominal significance level $\alpha = 0.05$ were considered. The studies used Fisher's combining function to determine the overall *ASL* value. The results of the simulations carried out to determine the size and power of the tests are presented in Tables 2–7. Estimated probabilities of rejection of the hypothesis H_0 when partial two–sided alternative hypotheses were taken under consideration are presented in Tables 2–4. In the case of partial one–sided alternative hypotheses, estimated probabilities are presented in Tables 5–7, respectively.

Table 6. Power estimates when $\mu_1 - \mu_2 = 0$ and $\sigma_1 / \sigma_2 > 1$, $\alpha = 0.05$, for samples $n_1 = 10, n_2 = 15$, (one–sided alternative hypotheses)

Model	Distribution parameters (μ_1, σ_1)							
	$(0,1.2)$	$(0,1.4)$	$(0,1.6)$	$(0,1.8)$	$(0,2)$	$(0,2.2)$	$(0,2.4)$	$(0,2.6)$
1	0.108	0.183	0.320	0.437	0.536	0.616	0.708	0.779
2	0.107	0.171	0.278	0.404	0.462	0.572	0.628	0.705
3	0.102	0.180	0.277	0.398	0.464	0.574	0.627	0.702
4	0.100	0.171	0.272	0.373	0.441	0.509	0.584	0.674
5	0.110	0.196	0.319	0.435	0.528	0.599	0.659	0.737

Source: Own calculation in R program.

Table 7. Power estimates when $\mu_1 - \mu_2 > 0$ and $\sigma_1 / \sigma_2 > 1$, $\alpha = 0.05$, for samples $n_1 = 10, n_2 = 15$, (one–sided alternative hypotheses)

Model	Distribution parameters (μ_1, σ_1)							
	$(0.2,1.2)$	$(0.4,1.4)$	$(0.6,1.6)$	$(0.8,1.8)$	$(1,2)$	$(1.2,2.2)$	$(1.4,2.4)$	$(1.6,2.6)$
1	0.167	0.345	0.526	0.672	0.806	0.865	0.928	0.949
2	0.159	0.322	0.490	0.625	0.745	0.838	0.895	0.933
3	0.158	0.320	0.487	0.616	0.748	0.846	0.894	0.924
4	0.156	0.313	0.473	0.600	0.745	0.811	0.881	0.906
5	0.160	0.352	0.518	0.661	0.801	0.866	0.931	0.948

Source: Own calculation in R program.

The size of the tests is shown in Tables 2 and 5 in the first column. For all models, the obtained estimated probabilities are close to the nominal level of significance $\alpha = 0.05$. The tests considered achieved comparable results in the case of small samples, the size of which was not equal. The tests used in models 1, 3 and 5 were the most powerful. The probabilities of detecting differences between populations increased with increasing differences between the respective location or scale parameters for both considered partial two– and one–sided alternative hypotheses.

5. Conclusions

The simulation research aimed to determine the ability of the presented location and scale tests to maintain the nominal probability of committing the type I error and the ability to obtain a high probability of rejecting a false zero hypothesis in the conditions of changing distribution parameters in populations from which samples were taken.

The tests that verify the hypothesis about the identity of location and scale parameters in the studied populations are presented. The article considered various forms of test statistics. A simulation study to determine the size and power of tests was carried out using permutation tests.

When analysing the results obtained it can be seen that the inference about the significance of differences between populations is possible with the use of the proposed solution. All testing procedures (under normality) ensured control of type I error at the assumed level of significance. The simulation analysis indicated that the proposed tests allowed the inference about the differences in location or scale parameters, as well as differences in both location and scale parameters of distributions. The results for the permutation Lepage test and permutation Cucconi test are also presented where two–sided alternative hypothesis is considered. Higher power of tests was achieved thanks to the use of a nonparametric procedure that uses Fisher's combining functions to evaluate the overall *ASL* value. The observed assessments of the probability of rejection of the null hypothesis were similar for various pairs of test statistics considered in the simulations. One advantage of the procedure presented in the article is also the possibility of formulating an alternative hypothesis in the form of partial directional hypotheses. The method can be used even in the case of small sample sizes. In the research, other forms of combining functions can be considered and a simulation study taking into account the various distributions of the studied variables can be performed. The direction of further research also concerns the extension of the method to a multidimensional case.

REFERENCES

ANDERSON, M. J., WALSH, D. C. I., CLARKE, K.R., GORLEY, R. N., GUERRA–CASTRO, E., (2017). Permutational Multivariate Analysis of Variance (PERMANOVA), Wiley StatsRef: Statistics Reference Online, pp. 1–15.

ANSARI, A. R., BRADLEY, R. A., (1960). Rank–sum tests for dispersions. Annals of Mathematical Statistics 31, pp. 1174–1189.

BALAKRISHNAN, N., MA, C. W., (1990). A comparative study of various tests for the equality of two population variances, Journal of Statistical Computation and Simulation, 35, pp. 41–89.

BONNINI, S., CORAIN, L., MAROZZI, M., SALMASO, L., (2014). Nonparametric Hypothesis Testing Rank and Permutation Methods with Applications in R, John Wiley & Sons, Ltd.

CHANG, C.–H., PAL, N., (2008). A Revisit to the Behrens–Fisher Problem: Comparison of Five Test Methods. Communications in Statistics – Simulation and Computation, 37, (6), pp. 1064–1085.

CONOVER, W. J., JOHNSON, M. E., JOHNSON, M. M., (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data, Technometrics, 23, pp. 351–361.

CUCCONI, O., (1968). Un nuovo test non parametrico per it confronto tra due gruppi campionori. Giornale degli Economisti, XXVII, pp. 225–248.

DURAN, B. S., TSAI, W. S., LEWIS, T. O., (1976). A class of location-scale tests, Biometrika, 63, pp. 173–176.

FISHER, R. A., (1932). Statistical Methods for Research Workers, 4 ed., Edinburgh: Oliver & Boyd.

GENG, S., WANG, W. J., MILLER, C., (1979). Small sample size comparisons of tests for homogeneity of variances by Monte-Carlo. Communications in Statistics – Simulation and Computation, 8, pp. 379–389.

GOGOI, P., GOGOI, B., (2017). Some Tests Procedures for Scale Differences. International Advanced Research Journal in Science, Engineering and Technology, Vol. 4, Issue 11, pp. 155–166.

HALL, I. J., (1972). Some comparisons of tests for equality of variances, Journal of Statistical Computation and Simulation, 1, pp. 183–194.

JANSSEN, A., PAULS, T., (2005). A Monte Carlo comparison of studentized bootstrap and permutation tests for heteroscedastic two–sample problems. Computational Statistics, 20 (3), pp. 369–383.

KESELMAN, H. J., GAMES, P. A., CLINCH, J. J., (1979). Tests for homogeneity of variance. Communications in Statistics – Simulation and Computation, 8, pp. 113–119.

KOŃCZAK, G., (2016). Testy permutacyjne, Teoria i zastosowania, Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.

LEPAGE, Y., (1971). A combination of Wilcoxon's and Ansari–Bradley's statistics, Biometrika 58, pp. 213–217.

LEPAGE, Y., (1973). A table for a combined Wilcoxon Ansari–Bradley statistic, Biometrika 60, pp. 113–116.

LIM, T.S., LOH, W. Y., (1996). A comparison of tests of equality of variances, Computational Statistics and Data Analysis, 22, pp. 287–301.

LIPTAK, I., (1958). On the combination of independent tests. Magyar Tudomanyos Akademia Matematikai Kutato Intezenek Kozlomenyei 3, pp. 127–141.

LUNDE, A., TIMMERMANN, A., (2004). Duration dependence in stock prices: an analysis of bull and bear markets, Journal of Business and Economic Statistics, 22, pp. 253–273.

MANN, H., WHITNEY, D., (1947). On a test of whether one of two random variables is stochastically larger than the other, Annals of Mathematical Statistics, 18, (1), pp. 50–60.

MAROZZI, M., (2004). A bi-aspect nonparametric test for the two-sample location problem, Computational Statistics and Data Analysis, 44, pp. 639–648.

MAROZZI, M., (2007). Multivariate tri–aspect non-parametric testing, Journal of Nonparametric Statistics, 19, pp. 269–282.

MAROZZI, M., (2008). The Lepage location–scale test revisited, Far East Journal of Theoretical Statistics 24, pp. 137–155.

MAROZZI, M., (2009). Some notes on the location–scale Cucconi test, Journal of Nonparametric Statistics, 21, 5, pp. 629–647.

MAROZZI, M., (2011). Levene type tests for the ratio of two scales, Journal of Statistical Computation and Simulation, 81, pp. 815–826.

MAROZZI, M., (2012a). A distribution free test for the equality of scales. Communication in Statistics – Simulation and Computation, 41, pp. 878–889.

MAROZZI, M., (2012b). A combined test for differences in scale based on the interquantile range. Statistical Papers, 53, pp. 61–72.

MOOD, A. M., (1954). On the asymptotic efficiency of certain nonparametric two–sample tests. Ann Math Stat 25, pp. 514–522.

MUCCIOLI, C., BELFORD, R., PODGOR, M., SAMPAIO, P., DE SMET, M., NUSSENBLATT, R., (1996). The diagnosis of intraocular inflammation and cytomegalovirus retinitis in HIV–infected patients by laser flare photometry, Ocular Immunology and Inflammation, 4, pp. 75–81.

MURAKAMI, H., (2007). Lepage type statistic based on the modified Baumgartner statistic, Computational Statistics & Data Analysis, 51, pp. 5061–5067.

NEUHAUSER, M., LEUCHS, A.-K., BALL, D., (2011). A new location-scale test based on a combination of the ideas of Levene and Lepage, Biometrical Journal, 53, pp. 525–534.

O'BRIEN, R. G., (1979). A general ANOVA method for robust test of additive models for variance, Journal of the American Statistical Association, 74, pp. 877–880.

PARK, H-I., (2015a). Simultaneous Tests with Combining Functions under Normality, Communications for Statistical Applications and Methods, Vol. 22, No. 6, pp. 639–646.

PARK, H-I., (2015b). Nonparametric Simultaneous Test Procedures, Revista Colombiana de Estadística, 38(1), pp. 107–121.

PESARIN, F., (2001). Multivariate Permutation Test with Applications in Biostatistics, Chichester: Wiley.

SALMASO, L., SOLARI, A., (2005). Multiple aspect testing for case-control designs, Metrika, 62, pp. 331–340.

TIPPETT, L. H. C., (1931). The Methods of Statistics, London: Williams and Norgate.

WILCOXON, F., (1949). Some rapid approximate statistical procedures, Stamford, CT: Stamford Research Laboratories, American Cyanamid Corporation.

YONETANI, T., GORDON, H. B., (2001). Abrupt changes as indicators of decadal climate variability, Climate Dynamics, 17, pp. 249–258.