

Nonparametric Simultaneous Test Procedures

Procedimientos de pruebas simultáneas no paramétricas

HYO-IL PARK^a

DEPARTMENT OF STATISTICS, CHONGJU UNIVERSITY, CHONGJU, SOUTH KOREA

Abstract

In this research we propose several nonparametric simultaneous test procedures for location and scale parameters. We construct test statistics based on linear rank statistics choosing a suitable combining function. We obtain the overall p -values by applying the permutation principle. We compare the efficiency amongst combining functions by obtaining empirical powers through a simulation study. We discuss some interesting aspects of our procedure as concluding remarks.

Key words: Combining function, Multi-aspect Test, Permutation Principle, Two-sample Location-scale Problem.

Resumen

En este artículo se propone un procedimiento de pruebas simultáneas no paramétricas para los parámetros de localización y escala. Se construyen los estadísticos de prueba basados en los estadísticos de rangos lineales para las subhipótesis nulas con la escogencia de una adecuada función de combinación; se obtienen los valores p al aplicar el principio de permutación; se compara la eficiencia entre las funciones de combinación mediante la obtención de las potencias empíricas a través de un estudio de simulación y por último se discuten algunos aspectos interesantes del procedimiento como conclusiones.

Palabras clave: funciones de combinación, parámetro de escala, parámetro de localización, principio de permutación, problema de dos muestras, pruebas simultáneas.

^aProfessor. E-mail: hipark@cju.ac.kr

1. Introduction

In order to improve test performance, some nonparametric testing procedures have adopted the trend of using several nonparametric test statistics simultaneously. One of the well-known procedures may be the versatile test (see Fleming, Harrington & O'Sullivan 1987), which combines several tests under the identical hypothesis with a suitable combining function and obtains overall p -values. Park (2011) considered several versatile tests with a group of quantile test statistics. On the other hand, one may consider reducing the scope of the null hypothesis by splitting it into several sub-null hypotheses according to the interesting aspects of the underlying distribution and then intersecting them. Pesarin (2001) has initiated this approach for nonparametric testing problems and named the multi-aspect test. Taking this approach, Marozzi (2004) combined the permutation t and median tests and Marozzi (2007) added also the Wilcoxon test in the combination to address the two-sample location problem. Also Salmaso & Solari (2005) considered the multi-aspect test for the case-control study. Brombin, Salmaso, Ferronato & Galzignato (2011) considered applying this multi-aspect test to the case of bio-medical data. Furthermore, the multi-aspect test has been very useful in addressing the scale problem (see Marozzi 2012*a*, 2012*b*, 2012*c*, 2012*d*).

Simultaneous use of several nonparametric tests can be applied also to the problem of testing location-scale parameters concurrently, for the two-sample case. In order to discuss this approach more concretely, let F_1 and F_2 be the distribution functions of the populations underlying the samples. Also let δ and η be the location translation and scale parameters, respectively. We assume that F_1 and F_2 satisfy the following location-scale model such that for all $x \in (-\infty, \infty)$,

$$F_2(x) = F_1\left(\frac{x - \delta}{\eta}\right), \quad (1)$$

for some $\delta \in (-\infty, \infty)$ and $\eta \in (0, \infty)$. In view of the simultaneous tests for δ and η under the model (1.1), we can express the null hypothesis as follows:

$$H_0 : \{\delta = 0\} \cap \{\eta = 1\}. \quad (2)$$

Then performing some reasonable nonparametric tests for each individual sub-null hypothesis and combining their results with a chosen combining function, one may obtain an overall p -value from the null distribution of the combined test statistic. This can be called a nonparametric simultaneous test procedure for the location-scale problem. Lepage (1971) has initiated this topic by combining the Wilcoxon rank sum and Ansari-Bradley's (Ansari & Bradley 1960) statistics for location and scale parameters, respectively using the quadratic form for the combining function. Lepage (1973) reported the exact critical points and significance levels for some selected sample sizes. Duran, Tsai & Lewis (1976) derived the asymptotic relative efficiency of Lepage's test with respect to another simultaneous test using Mood's statistic, for the scale parameter (Mood 1954). Lepage's procedure has been reviewed and discussed extensively Podgor & Gastwirth (1994), Zhang (2006) and modified in various ways (Murakami 2007, Rublik 2009, Neuhäuser, Leuchs & Ball 2011, Marozzi 2012*a*). Marozzi (2013) reviewed and compared these tests by obtaining empirical powers through an extensive simulation study. We note that

all the reviewed simultaneous test statistics have adopted the quadratic form to combine two individual test statistics for the sub-null hypotheses while those by Marozzi (2012a) have followed the combining function approach. A rather different test is the Cucconi one (Marozzi 2009) which is not a quadratic form combining a test statistic for location and another for scale but considers squared ranks and squared contrary-ranks.

For obtaining an overall p -value to complete a simultaneous test, one has to derive the null distribution of the chosen combining function. For this purpose, one may derive the asymptotic normality with the large sample approximation theorem. However nowadays, it is common to apply the permutation principle (Good 2000), which uses the re-sampling method. The permutation principle has been proposed by Fisher (1932) yet only until quite recently with the rapid development of computer capability has begun to be widely used. In passing, we note that the permutation principle yields an exact test (Good 2000).

In this research, we propose a nonparametric simultaneous test procedure for (2) with several combining functions for the two sample problem. We particularly note that we have excluded the use of the quadratic form for the combining function to accommodate various types of alternatives. The rest of this paper will be organized as follows. In Section 2, we construct the test statistics for the simultaneous test through p -values from individual partial tests for the sub-null hypotheses, $H_{01} : \delta = 0$ and $H_{02} : \eta = 1$, using the (1) Fisher (1932), (2) Liptak (1958) and (3) Tippett (1931) combining functions. Then we obtained overall p -values for any chosen combining function by applying the permutation principle. In Section 3, we investigate the performance of our procedure and compare it with other tests by obtaining empirical powers through a simulation study. For this, we consider the following two cases separately: both sub- alternatives are one- and two-sided. We also considered the Wilcoxon rank sum test for location and Mood's and Ansari-Bradley's tests for scale to investigate our procedure. Then we discuss some interesting features of the nonparametric simultaneous tests, the use of the limiting distributions for the combining functions and the bootstrap method for obtaining overall p -values as concluding remarks in Section 4.

2. Formulation of Nonparametric Simultaneous Tests

Let X_{11}, \dots, X_{1n_1} and X_{21}, \dots, X_{2n_2} be two independent random samples from populations with distribution functions F_1 and F_2 , respectively. We assume that F_i is unknown but continuous for each i , $i = 1, 2$ satisfying the relation (1). We are interested in testing (2) which requires a simultaneous test procedure for both location and scale parameters. For this purpose, we will use the nonparametric multi-aspect testing approach. This means that first of all we choose suitable linear rank tests for the sub-null hypotheses $H_{01} : \delta = 0$ and $H_{02} : \eta = 1$ and then combine the results of the two individual tests to obtain an overall p -value. Randles & Wolfe (1979) have studied and summarized extensively linear rank

tests for testing hypothesis on location and scale parameters for the two-sample problems. Let L and S be the respective linear rank statistics for testing the two sub-null hypotheses $H_{01} : \delta = 0$ and $H_{02} : \eta = 1$ such as

$$L = \sum_{j=1}^{n_1} a(R_{1j}) \quad \text{and} \quad S = \sum_{j=1}^{n_1} b(R_{1j}),$$

where a and b are the score functions and R_{1j} , the rank of X_{1j} from the combined sample. For example, one may choose the Wilcoxon rank sum statistic for L and Mood's statistic for S . Without loss of generality, we assume that both statistics L and S are of the standardized forms. In order to proceed our discussion further for the construction of the test statistics for testing (2), let λ_1 and λ_2 be the p -values for testing the sub-null hypotheses, $H_{01} : \delta = 0$ and $H_{02} : \eta = 1$ based on the statistics L and S , respectively. Then for obtaining an overall p -value for simultaneously testing (2), we consider a suitable combining function to combine the two individual p -values, λ_1 and λ_2 . For this matter, Pesarin (2001) has in detail reviewed and summarized several useful combining functions. In the following, we introduce three types of combining functions, which will be used in the simulation study.

(1) The Fisher omnibus combining function (Fisher 1932) corresponds to

$$C_F = -2\{\log(\lambda_1) + \log(\lambda_2)\}.$$

We note that if two partial test statistics are independent and continuous, then asymptotically C_F follows a χ^2 distribution with 4 degrees of freedom under (1.1).

The Liptak combining function (Liptak 1958) corresponds to

$$C_L = \Phi^{-1}(1 - \lambda_1) + \Phi^{-1}(1 - \lambda_2),$$

where Φ^{-1} is the inverse of the standard normal distribution function.

(3) The Tippett combining function (Tippett 1931) corresponds to

$$C_T = \max\{1 - \lambda_1, 1 - \lambda_2\}.$$

To complete the simultaneous test for (2), we have to obtain the null distribution of any chosen combining function to compute an overall p -value. We may achieve this task by applying the permutation principle. Even though the permutation principle yields an exact test, the excessive computational burden for the consideration of all the permutational configurations leads us to take the Monte-Carlo approach for the re-sampling phase, which yields an approximate result.

In the next section, we perform a simulation study to compare the efficiency among the tests corresponding to the combining functions listed above with other well-known tests. With Fisher, Liptak and Tippett tests in the sequel, we indicate those using the combining functions (1), (2) and (3), respectively. In addition, we consider the Lepage (1971), Pettitt (1976) and Neuhäuser et al. (2011) tests for this comparison study. We note that the Lepage, Pettitt and Neuhäuser tests take the quadratic form for their test statistics. This implies that these three tests cannot be applied to the one-sided alternative for each individual test whereas ours can. For this reason, we will carry out the simulation study in two parts under the schemes of the one- and two-sided alternatives separately.

3. A Simulation Study

In this section, we investigate the performance of the tests by estimating their empirical powers through a simulation study. First of all, for the one-sided alternative, we compare the empirical powers among our proposed tests. Then we consider the Lepage, Pettitt and Neuhäuser tests with ours for the two-sided alternative. We consider the following six different distributions: normal, Cauchy, half-Cauchy, exponential, uniform and double exponential with unit variance except the Cauchy and half-Cauchy distributions. For our procedure with Lepage test, we use the Wilcoxon test for the location and the Mood and Ansari-Bradley tests for the scale. In view of the null hypothesis (2), we considered values of the pair (δ, η) varying from (0,1) to (1,2) with the increment 0.2 for each parameter. We note that (0,1) are the values of the pair (δ, η) under the null hypothesis, 2. The sample sizes for this study were chosen as (10,10), (10,20) and (20,10) and the nominal significance level is 0.05 for all cases. The simulation has been conducted with SAS/IML on the PC version and all the results in the tables are based on 10,000 simulations with the Monte-Carlo method and within a simulation, we applied the permutation principle by 5,000 iterations also with the Monte-Carlo approach to estimate the distribution for each test. During the revision of this paper, a referee brought a paper by Marozzi (2014) to our attention containing a discussion on the optimal choice of the number of iterations with the permutation principle given the Monte-Carlo simulation size. As suggested, one could say that 5,000 iterations in this work may be much higher than the recommended numbers. The simulation results are all summarized in Tables 1 through 12 for the one-sided alternative and Tables 2 to 12 for the two-sided alternative.

We note that the two kinds of procedures based on Mood and Ansari-Bradley tests for the scale yielded almost identical results for both cases. For the one-sided alternative, the Liptak and Fisher tests show high performance for the normal, Cauchy, uniform and double exponential distributions while the Tippett one achieves better performance for the half-Cauchy and exponential distributions. For the two-sided alternative, the Tippett, Liptak and Fisher tests show high performance for the normal, Cauchy and double-exponential distributions while the Pettitt one obtains higher power than any other test for the exponential and uniform distributions. We note that the Neuhäuser test yields very high performance for the uniform distribution but does not achieve the nominal significance level for the half-Cauchy and exponential distributions which are skewed. Therefore, one should apply the Neuhäuser test with caution when the underlying distribution is skewed. The Lepage test shows poor performance for all cases. We likewise note that the Liptak and Fisher tests seem to be suitable for symmetric distributions while the Tippett test may be appropriate for the skewed ones for the one-sided alternative.

In general, the Lepage and Pettitt tests appear to be somewhat conservative since their empirical type I errors are lower than the nominal significance level 0.05, while our tests tend to achieve those slightly higher. However, the phenomena tests appearing conservative or not seem not to affect the power of same. Finally we note that powers may depend on whether the two sample sizes are equal or not.

TABLE 1: Power estimates for normal distribution (one-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.055	0.127	0.220	0.328	0.429	0.526
		20,10	0.051	0.153	0.301	0.450	0.582	0.695
		10,20	0.049	0.125	0.232	0.352	0.475	0.589
	Liptak	10,10	0.053	0.152	0.286	0.441	0.575	0.686
		20,10	0.057	0.188	0.380	0.563	0.709	0.803
		10,20	0.046	0.149	0.312	0.502	0.672	0.796
	Fisher	10,10	0.052	0.144	0.273	0.420	0.550	0.663
		20,10	0.054	0.181	0.363	0.546	0.690	0.795
		10,20	0.046	0.144	0.294	0.471	0.643	0.767
Wilcoxon and Ansari-Bradley	Tippett	10,10	0.059	0.132	0.221	0.320	0.422	0.515
		20,10	0.053	0.151	0.292	0.429	0.558	0.665
		10,20	0.052	0.128	0.231	0.349	0.465	0.577
	Liptak	10,10	0.060	0.156	0.289	0.433	0.564	0.671
		20,10	0.060	0.187	0.368	0.546	0.680	0.775
		10,20	0.049	0.153	0.312	0.488	0.648	0.771
	Fisher	10,10	0.057	0.153	0.275	0.414	0.537	0.649
		20,10	0.058	0.181	0.351	0.526	0.663	0.760
		10,20	0.049	0.147	0.289	0.458	0.621	0.743

TABLE 2: Power estimates for normal distribution (two-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.054	0.077	0.138	0.211	0.288	0.363
		20,10	0.051	0.098	0.209	0.340	0.461	0.580
		10,20	0.052	0.066	0.122	0.205	0.289	0.375
	Liptak	10,10	0.042	0.060	0.116	0.192	0.290	0.384
		20,10	0.046	0.098	0.208	0.359	0.505	0.628
		10,20	0.045	0.052	0.104	0.185	0.292	0.412
	Fisher	10,10	0.048	0.069	0.133	0.222	0.320	0.421
		20,10	0.045	0.103	0.222	0.375	0.521	0.644
		10,20	0.047	0.060	0.117	0.210	0.325	0.446
	Lepage	10,10	0.039	0.057	0.113	0.188	0.279	0.370
		20,10	0.039	0.093	0.209	0.356	0.500	0.624
		10,20	0.043	0.052	0.108	0.192	0.296	0.410
Wilcoxon and Ansari-Bradley	Tippett	10,10	0.060	0.084	0.144	0.216	0.289	0.365
		20,10	0.056	0.097	0.190	0.313	0.426	0.532
		10,20	0.054	0.074	0.136	0.221	0.306	0.393
	Liptak	10,10	0.052	0.073	0.132	0.208	0.305	0.399
		20,10	0.052	0.099	0.199	0.335	0.477	0.591
		10,20	0.051	0.064	0.120	0.209	0.316	0.434
	Fisher	10,10	0.058	0.081	0.144	0.230	0.327	0.423
		20,10	0.052	0.099	0.210	0.348	0.481	0.600
		10,20	0.053	0.067	0.136	0.232	0.346	0.465
	Lepage	10,10	0.039	0.061	0.114	0.184	0.268	0.360
		20,10	0.043	0.086	0.186	0.318	0.447	0.564
		10,20	0.045	0.056	0.117	0.201	0.306	0.413
Pettitt	10,10	0.042	0.059	0.100	0.153	0.218	0.278	
	20,10	0.047	0.082	0.162	0.268	0.375	0.482	
	10,20	0.045	0.054	0.097	0.165	0.246	0.346	
Neuhäuser	10,10	0.034	0.055	0.106	0.180	0.265	0.354	
	20,10	0.036	0.061	0.139	0.250	0.373	0.497	
	10,20	0.042	0.067	0.143	0.254	0.388	0.523	

TABLE 3: Power estimates for Cauchy distribution (one-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.052	0.089	0.136	0.184	0.233	0.278
		20,10	0.050	0.094	0.153	0.221	0.296	0.368
		10,20	0.051	0.092	0.145	0.201	0.253	0.307
	Liptak	10,10	0.055	0.102	0.159	0.223	0.284	0.342
		20,10	0.058	0.117	0.192	0.278	0.364	0.445
		10,20	0.049	0.104	0.176	0.253	0.324	0.398
	Fisher	10,10	0.056	0.099	0.154	0.214	0.276	0.339
		20,10	0.052	0.112	0.181	0.265	0.355	0.437
		10,20	0.048	0.100	0.169	0.245	0.311	0.381
Wilcoxon and Ansari-Bradley	Tippett	10,10	0.056	0.094	0.146	0.195	0.243	0.289
		20,10	0.050	0.098	0.163	0.236	0.311	0.383
		10,20	0.053	0.096	0.148	0.210	0.263	0.315
	Liptak	10,10	0.061	0.111	0.173	0.236	0.298	0.365
		20,10	0.061	0.123	0.203	0.293	0.381	0.467
		10,20	0.052	0.109	0.182	0.261	0.336	0.406
	Fisher	10,10	0.061	0.104	0.165	0.231	0.294	0.356
		20,10	0.056	0.115	0.193	0.282	0.372	0.456
		10,20	0.051	0.106	0.176	0.254	0.325	0.391

TABLE 4: Power estimates for Cauchy distribution (two-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.054	0.060	0.083	0.116	0.146	0.181
		20,10	0.051	0.068	0.100	0.149	0.205	0.264
		10,20	0.049	0.056	0.079	0.112	0.144	0.182
	Liptak	10,10	0.044	0.050	0.070	0.100	0.136	0.172
		20,10	0.044	0.064	0.102	0.151	0.208	0.267
		10,20	0.046	0.046	0.066	0.096	0.135	0.178
	Fisher	10,10	0.049	0.058	0.079	0.112	0.150	0.185
		20,10	0.046	0.067	0.106	0.159	0.219	0.285
		10,20	0.046	0.049	0.074	0.107	0.149	0.196
	Lepage	10,10	0.040	0.047	0.066	0.095	0.129	0.164
		20,10	0.041	0.060	0.096	0.147	0.205	0.268
		10,20	0.040	0.044	0.066	0.097	0.136	0.181
Wilcoxon and Ansari-Bradley	Tippett	10,10	0.060	0.068	0.090	0.126	0.159	0.198
		20,10	0.053	0.066	0.101	0.151	0.206	0.267
		10,20	0.053	0.061	0.092	0.130	0.166	0.202
	Liptak	10,10	0.055	0.064	0.084	0.116	0.150	0.188
		20,10	0.047	0.069	0.107	0.158	0.218	0.280
		10,20	0.052	0.056	0.079	0.113	0.156	0.197
	Fisher	10,10	0.059	0.067	0.089	0.125	0.164	0.205
		20,10	0.050	0.071	0.107	0.164	0.227	0.293
		10,20	0.051	0.060	0.088	0.125	0.168	0.216
	Lepage	10,10	0.042	0.050	0.068	0.098	0.130	0.168
		20,10	0.042	0.059	0.093	0.145	0.202	0.266
		10,20	0.043	0.049	0.075	0.108	0.148	0.193
Pettitt	10,10	0.043	0.050	0.067	0.090	0.111	0.135	
	20,10	0.043	0.057	0.085	0.122	0.164	0.208	
	10,20	0.046	0.050	0.066	0.093	0.124	0.154	
Neuhäuser	10,10	0.051	0.059	0.077	0.105	0.136	0.170	
	20,10	0.049	0.063	0.091	0.130	0.179	0.229	
	10,20	0.052	0.060	0.091	0.127	0.176	0.230	

TABLE 5: Power estimates for half-Cauchy distribution (one-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.051	0.110	0.239	0.384	0.518	0.628
		20,10	0.048	0.125	0.295	0.492	0.668	0.786
		10,20	0.050	0.140	0.316	0.495	0.642	0.747
	Liptak	10,10	0.055	0.070	0.099	0.141	0.184	0.251
		20,10	0.056	0.086	0.145	0.228	0.333	0.432
		10,20	0.048	0.051	0.068	0.102	0.149	0.210
	Fisher	10,10	0.053	0.085	0.168	0.274	0.387	0.490
		20,10	0.052	0.101	0.217	0.370	0.527	0.658
		10,20	0.050	0.097	0.210	0.359	0.496	0.611
Wilcoxon and Ansari- Bradley	Tippett	10,10	0.055	0.112	0.239	0.384	0.518	0.628
		20,10	0.050	0.120	0.295	0.492	0.668	0.786
		10,20	0.053	0.149	0.317	0.495	0.642	0.747
	Liptak	10,10	0.058	0.082	0.114	0.152	0.197	0.264
		20,10	0.059	0.096	0.151	0.232	0.325	0.418
		10,20	0.051	0.066	0.089	0.125	0.173	0.230
	Fisher	10,10	0.058	0.094	0.178	0.283	0.394	0.502
		20,10	0.055	0.108	0.222	0.372	0.527	0.656
		10,20	0.053	0.106	0.222	0.370	0.502	0.616

TABLE 6: Power estimates for half-Cauchy distribution (two-sided).

Test	(n_1, n_2)	(δ, η)							
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0		
Wilcoxon and Mood	Tippett	10,10	0.053	0.101	0.216	0.350	0.475	0.587	
		20,10	0.052	0.104	0.246	0.421	0.589	0.718	
		10,20	0.051	0.147	0.336	0.524	0.697	0.788	
	Liptak	10,10	0.047	0.081	0.185	0.326	0.453	0.570	
		20,10	0.046	0.070	0.177	0.325	0.474	0.576	
		10,20	0.046	0.154	0.383	0.604	0.760	0.865	
	Fisher	10,10	0.050	0.091	0.215	0.379	0.527	0.657	
		20,10	0.047	0.089	0.224	0.423	0.624	0.777	
		10,20	0.047	0.160	0.386	0.599	0.757	0.863	
	Lepage	10,10	0.041	0.078	0.185	0.337	0.477	0.604	
		20,10	0.042	0.081	0.208	0.396	0.590	0.747	
		10,20	0.042	0.149	0.363	0.574	0.738	0.845	
	Wilcoxon and Ansari- Bradley	Tippett	10,10	0.059	0.104	0.224	0.360	0.491	0.607
			20,10	0.056	0.111	0.258	0.438	0.612	0.736
			10,20	0.054	0.139	0.318	0.505	0.659	0.770
Liptak		10,10	0.057	0.096	0.222	0.379	0.523	0.647	
		20,10	0.051	0.087	0.215	0.389	0.553	0.666	
		10,20	0.052	0.151	0.356	0.565	0.726	0.835	
Fisher		10,10	0.059	0.105	0.239	0.410	0.561	0.687	
		20,10	0.054	0.102	0.255	0.473	0.674	0.817	
		10,20	0.053	0.155	0.364	0.569	0.728	0.837	
Lepage		10,10	0.042	0.079	0.189	0.343	0.489	0.620	
		20,10	0.045	0.089	0.223	0.425	0.628	0.779	
		10,20	0.044	0.136	0.335	0.538	0.700	0.815	
Pettitt		10,10	0.042	0.085	0.211	0.362	0.509	0.628	
		20,10	0.046	0.103	0.260	0.471	0.665	0.799	
		10,20	0.042	0.140	0.344	0.552	0.714	0.823	
Neuhäuser	10,10	0.132	0.166	0.247	0.346	0.444	0.533		
	20,10	0.137	0.178	0.284	0.423	0.561	0.673		
	10,20	0.132	0.195	0.322	0.450	0.567	0.658		

TABLE 7: Power estimates for exponential distribution (one-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.051	0.167	0.420	0.665	0.836	0.923
		20,10	0.052	0.194	0.530	0.808	0.938	0.983
		10,20	0.053	0.227	0.532	0.782	0.910	0.968
	Liptak	10,10	0.052	0.101	0.197	0.347	0.508	0.651
		20,10	0.054	0.142	0.326	0.565	0.766	0.888
		10,20	0.046	0.075	0.141	0.260	0.406	0.557
Fisher	10,10	0.053	0.131	0.321	0.552	0.741	0.861	
	20,10	0.053	0.163	0.437	0.718	0.886	0.961	
	10,20	0.047	0.162	0.408	0.662	0.838	0.928	
Wilcoxon and Ansari- Bradley	Tippett	10,10	0.053	0.169	0.421	0.665	0.836	0.923
		20,10	0.056	0.197	0.530	0.808	0.938	0.983
		10,20	0.055	0.229	0.533	0.783	0.910	0.968
	Liptak	10,10	0.059	0.119	0.214	0.354	0.495	0.624
		20,10	0.057	0.152	0.319	0.538	0.731	0.863
		10,20	0.051	0.100	0.187	0.307	0.451	0.595
Fisher	10,10	0.056	0.143	0.332	0.560	0.742	0.861	
	20,10	0.057	0.173	0.436	0.713	0.881	0.959	
	10,20	0.052	0.177	0.426	0.673	0.844	0.931	

TABLE 8: Power estimates for exponential distribution (two-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.052	0.131	0.333	0.566	0.757	0.874
		20,10	0.052	0.143	0.423	0.719	0.894	0.968
		10,20	0.048	0.193	0.471	0.723	0.875	0.949
	Liptak	10,10	0.047	0.094	0.224	0.369	0.506	0.611
		20,10	0.045	0.091	0.243	0.424	0.595	0.750
		10,20	0.045	0.200	0.496	0.745	0.884	0.954
	Fisher	10,10	0.050	0.119	0.312	0.547	0.741	0.866
		20,10	0.049	0.117	0.366	0.667	0.870	0.959
		10,20	0.046	0.210	0.511	0.762	0.900	0.962
	Lepage	10,10	0.040	0.098	0.277	0.509	0.709	0.844
		20,10	0.044	0.110	0.354	0.658	0.867	0.955
		10,20	0.041	0.193	0.489	0.745	0.890	0.958
Wilcoxon and Ansari- Bradley	Tippett	10,10	0.056	0.136	0.341	0.576	0.763	0.879
		20,10	0.054	0.150	0.434	0.728	0.897	0.970
		10,20	0.054	0.189	0.463	0.721	0.874	0.948
	Liptak	10,10	0.056	0.124	0.297	0.491	0.655	0.770
		20,10	0.053	0.113	0.309	0.513	0.675	0.792
		10,20	0.050	0.194	0.471	0.714	0.860	0.936
	Fisher	10,10	0.060	0.139	0.351	0.591	0.781	0.892
		20,10	0.054	0.137	0.412	0.718	0.900	0.971
		10,20	0.052	0.206	0.498	0.751	0.893	0.958
	Lepage	10,10	0.043	0.107	0.294	0.530	0.731	0.857
		20,10	0.046	0.120	0.381	0.693	0.888	0.967
		10,20	0.043	0.184	0.469	0.728	0.881	0.952
Pettitt	10,10	0.040	0.135	0.377	0.624	0.812	0.912	
	20,10	0.044	0.162	0.488	0.784	0.932	0.982	
	10,20	0.042	0.221	0.545	0.803	0.923	0.977	
Neuhäuser	10,10	0.083	0.160	0.335	0.535	0.708	0.831	
	20,10	0.095	0.182	0.418	0.679	0.861	0.949	
	10,20	0.091	0.218	0.446	0.670	0.829	0.921	

TABLE 9: Power estimates for uniform distribution (one-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.051	0.233	0.471	0.692	0.836	0.927
		20,10	0.052	0.298	0.595	0.799	0.926	0.973
		10,20	0.053	0.270	0.583	0.824	0.943	0.984
	Liptak	10,10	0.053	0.298	0.580	0.771	0.881	0.942
		20,10	0.057	0.409	0.724	0.884	0.954	0.982
		10,20	0.051	0.298	0.612	0.815	0.919	0.963
	Fisher	10,10	0.052	0.281	0.562	0.763	0.884	0.949
		20,10	0.054	0.381	0.698	0.870	0.951	0.982
		10,20	0.052	0.310	0.653	0.877	0.966	0.990
Wilcoxon and Ansari- Bradley	Tippett	10,10	0.053	0.238	0.486	0.708	0.845	0.931
		20,10	0.057	0.298	0.598	0.804	0.927	0.973
		10,20	0.056	0.274	0.601	0.839	0.950	0.987
	Liptak	10,10	0.060	0.312	0.598	0.783	0.889	0.943
		20,10	0.059	0.392	0.701	0.864	0.944	0.975
		10,20	0.054	0.325	0.670	0.869	0.953	0.982
	Fisher	10,10	0.057	0.294	0.580	0.785	0.898	0.954
		20,10	0.057	0.368	0.685	0.862	0.946	0.981
		10,20	0.055	0.329	0.694	0.903	0.976	0.994

TABLE 10: Power estimates for uniform distribution (two-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.052	0.143	0.331	0.557	0.743	0.868
		20,10	0.052	0.203	0.465	0.701	0.862	0.948
		10,20	0.048	0.164	0.426	0.707	0.885	0.962
	Liptak	10,10	0.047	0.097	0.225	0.349	0.455	0.550
		20,10	0.045	0.220	0.511	0.730	0.861	0.936
		10,20	0.045	0.098	0.243	0.407	0.573	0.725
	Fisher	10,10	0.050	0.125	0.312	0.523	0.702	0.833
		20,10	0.049	0.226	0.516	0.736	0.879	0.951
		10,20	0.046	0.129	0.371	0.649	0.848	0.945
Lepage	10,10	0.040	0.107	0.279	0.489	0.678	0.816	
	20,10	0.044	0.209	0.492	0.720	0.869	0.948	
	10,20	0.041	0.124	0.357	0.640	0.845	0.944	
Wilcoxon and Ansari- Bradley	Tippett	10,10	0.056	0.149	0.346	0.572	0.753	0.873
		20,10	0.054	0.201	0.467	0.706	0.864	0.949
		10,20	0.054	0.173	0.441	0.720	0.891	0.964
	Liptak	10,10	0.056	0.133	0.317	0.495	0.640	0.743
		20,10	0.053	0.219	0.492	0.705	0.837	0.917
		10,20	0.050	0.124	0.330	0.528	0.674	0.783
	Fisher	10,10	0.060	0.148	0.368	0.593	0.765	0.872
		20,10	0.054	0.223	0.507	0.732	0.872	0.951
		10,20	0.052	0.151	0.436	0.729	0.896	0.967
Lepage	10,10	0.043	0.114	0.310	0.530	0.714	0.840	
	20,10	0.046	0.199	0.481	0.712	0.861	0.946	
	10,20	0.043	0.136	0.405	0.701	0.882	0.963	
Pettitt	10,10	0.040	0.145	0.379	0.628	0.804	0.910	
	20,10	0.044	0.228	0.562	0.802	0.929	0.977	
	10,20	0.042	0.179	0.497	0.787	0.933	0.982	
Neuhäuser	10,10	0.037	0.114	0.345	0.632	0.829	0.931	
	20,10	0.041	0.162	0.487	0.777	0.911	0.970	
	10,20	0.044	0.173	0.515	0.836	0.858	0.990	

TABLE 11: Power estimates for double exponential distribution (one-sided).

Test	(n_1, n_2)	(δ, η)						
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0	
Wilcoxon and Mood	Tippett	10,10	0.051	0.127	0.228	0.353	0.472	0.567
		20,10	0.052	0.155	0.300	0.467	0.621	0.741
		10,20	0.053	0.113	0.254	0.385	0.514	0.629
	Liptak	10,10	0.053	0.150	0.291	0.450	0.585	0.699
		20,10	0.057	0.185	0.383	0.587	0.746	0.849
		10,20	0.051	0.134	0.323	0.487	0.638	0.760
	Fisher	10,10	0.052	0.143	0.283	0.433	0.574	0.684
		20,10	0.054	0.180	0.367	0.571	0.728	0.835
		10,20	0.052	0.129	0.316	0.483	0.632	0.761
Wilcoxon and Ansari- Bradley	Tippett	10,10	0.053	0.130	0.234	0.357	0.475	0.571
		20,10	0.057	0.157	0.303	0.465	0.618	0.736
		10,20	0.056	0.115	0.253	0.382	0.509	0.624
	Liptak	10,10	0.060	0.157	0.297	0.454	0.587	0.693
		20,10	0.059	0.190	0.387	0.588	0.744	0.844
		10,20	0.054	0.137	0.316	0.473	0.609	0.726
	Fisher	10,10	0.057	0.149	0.287	0.440	0.576	0.684
		20,10	0.057	0.183	0.367	0.571	0.728	0.832
		10,20	0.055	0.132	0.310	0.473	0.616	0.740

TABLE 12: Power estimates for double exponential distribution (two-sided).

Test	(n_1, n_2)	(δ, η)							
		0.0,1.0	0.2,1.2	0.4,1.4	0.6,1.6	0.8,1.8	1.0,2.0		
Wilcoxon and Mood	Tippett	10,10	0.057	0.140	0.269	0.390	0.484	0.560	
		20,10	0.052	0.099	0.210	0.344	0.494	0.622	
		10,20	0.048	0.071	0.146	0.246	0.343	0.446	
	Liptak	10,10	0.049	0.120	0.267	0.417	0.538	0.635	
		20,10	0.045	0.091	0.207	0.366	0.533	0.678	
		10,20	0.045	0.060	0.130	0.228	0.335	0.442	
	Fisher	10,10	0.052	0.138	0.293	0.447	0.571	0.665	
		20,10	0.049	0.098	0.223	0.380	0.554	0.692	
		10,20	0.046	0.067	0.145	0.256	0.385	0.507	
	Lepage	10,10	0.043	0.118	0.256	0.396	0.524	0.619	
		20,10	0.044	0.090	0.209	0.359	0.533	0.670	
		10,20	0.041	0.059	0.130	0.237	0.361	0.478	
	Wilcoxon and Ansari- Bradley	Tippett	10,10	0.064	0.142	0.275	0.400	0.501	0.584
			20,10	0.054	0.098	0.206	0.336	0.481	0.607
			10,20	0.054	0.081	0.156	0.258	0.354	0.457
Liptak		10,10	0.058	0.131	0.286	0.435	0.558	0.649	
		20,10	0.053	0.097	0.213	0.364	0.528	0.670	
		10,20	0.050	0.071	0.142	0.232	0.338	0.442	
Fisher		10,10	0.063	0.146	0.304	0.461	0.583	0.676	
		20,10	0.054	0.101	0.223	0.373	0.542	0.679	
		10,20	0.052	0.076	0.158	0.269	0.390	0.505	
Lepage		10,10	0.043	0.113	0.247	0.388	0.510	0.608	
		20,10	0.046	0.086	0.198	0.342	0.510	0.646	
		10,20	0.043	0.065	0.134	0.243	0.355	0.468	
Pettitt		10,10	0.042	0.079	0.149	0.224	0.294	0.352	
		20,10	0.044	0.091	0.197	0.333	0.470	0.591	
		10,20	0.042	0.074	0.157	0.271	0.391	0.514	
Neuhäuser	10,10	0.082	0.100	0.133	0.194	0.266	0.341		
	20,10	0.042	0.078	0.166	0.285	0.421	0.547		
	10,20	0.043	0.077	0.168	0.293	0.429	0.562		

4. Some Concluding Remarks

Liptak (1958) studied the following statistic of type

$$\sum_{i=1}^k \alpha_i \Psi^{-1}(G_i(T_i)),$$

where Ψ^{-1} is the inverse of an arbitrary distribution function Ψ and α_i 's are arbitrary weights. Also T_i is the summarized test statistic for the i th parameter and G_i , the null distribution function of T_i , $i = 1, \dots, k$. The choice of Φ for Ψ allows to use the table of the standard normal distribution from the point of view of the distribution theory. Also the choice of $\alpha_1 = \alpha_2 = 1$ implies that we treat location and scale aspects as equally important in this study. In situations where an aspect is considered more important than the other, we can assess this by allocating more weight to the important one. For more information related to the Liptak combination function, you may refer to van Zwet & Oosterhoff (1967).

The simultaneous test for location and scale parameters can be regarded as one other application of the multi-aspect test based on the two parameters-location and scale when one considers the following null hypothesis $H_0 : F_1 = F_2$. However we do note that the multi-aspect test requires multiple partial tests for a parameter, while the simultaneous test selects a partial test for a sub-null hypothesis under the location-scale model (1.1). As an example, under the assumption that F_1 and F_2 may differ only in the location parameter with the same variance, one should use a multi-aspect test by applying the Wilcoxon and two-sample t -tests simultaneously. On the other hand, if one cannot assure which parameter may be different or assumes that both parameters may be different, it would be appropriate to apply a simultaneous test by using the Wilcoxon and Mood tests for the location and the scale parameters, respectively. Therefore the use of a simultaneous test would lead to loss of efficiency in terms of power to compared with a multi-aspect test when the location differs only between F_1 and F_2 and vice versa.

In order to obtain overall p -values in the simulation study, we have applied the permutation principle twice successively in loop style, thus demanding excessive computational work. To alleviate this computational burden, one may obtain p -values using the asymptotic normality approach based on the large sample approximation theory. For purposes of convenience in our discussion, we assume that L and S are of the standardized form under 2, which implies that the mean and variance are 0 and 1, respectively, for both cases. Then it is well-known that the limiting null distributions of L and S are standard normal (Randles & Hogg 1971). Also we note that L and S are uncorrelated when S is Ansari-Bradley or Mood's statistic (Duran et al. 1976). Then one may conclude that the limiting null distribution of $C_L = \Phi^{-1}(1 - \lambda_1) + \Phi^{-1}(1 - \lambda_2)$ is normal with mean 0 and variance 2. For the null distribution of C_T , you may refer to Pesarin & Salmaso (2010) for some detailed discussion. For Fisher's combining function, it is obvious that the limiting null distribution becomes a χ^2 distribution with 4 degrees of freedom.

However we have used p -values instead of linear rank statistics for the construction of test statistics in this study. This may require some additional computational

work but can be applied directly to any type of alternatives since each p -value should be obtained with respect to the corresponding sub-alternative. Therefore one can apply our procedure without any modification even when the two sub-alternatives are different such as two-sided for the location but one-sided for the scale parameter. In order to clarify our arguments, we will consider an example. Suppose that we apply the Wilcoxon rank sum test for testing $H_{01} : \delta = 0$ against $H_{11} : \delta \neq 0$ and the Mood test for testing $H_{02} : \eta = 1$ against $H_{12} : \eta > 1$. We can obtain the two p -values, λ_1 and λ_2 based on the given data using the distributions of L and S . For any chosen combining function C discussed in the previous section, we can obtain an overall p -value using the permutation principle to complete the test procedure for testing 2 against $H_1 : \{\delta \neq 0\} \cup \{\eta > 1\}$.

In this paper, we took the permutation framework. Another popular re-sampling framework is the bootstrap method (Efron 1979, Shao & Tu 1995). The difference between the bootstrap and permutation methods is being that the bootstrap method re-samples with replacement whereas the permutation method does so without replacement from the original pooled sample. However the difference can be significant for some cases (Good 2000). Bootstrap tests are more flexible than permutation tests because they may also be used when the null hypothesis is not a hypothesis of invariance and particularly in certain cases when the exchangeability condition is not satisfied. On the other hand, they are data-dependent without being strictly conditional procedures. In other words, for the finite sample sizes, inferential interpretations of bootstrap tests are not completely clear because they are neither conditional nor unconditional procedures. For a more detailed discussion, see Pesarin & Salmaso (2010).

Acknowledgement

The author wishes to express his sincere appreciation to the two anonymous referees for suggesting constructive opinions and pointing out errors.

[Received: December 2013 — Accepted: May 2014]

References

- Ansari, A. R. & Bradley, R. A. (1960), 'Rank-sum tests for dispersions', *Annals of Mathematical Statistics* **31**, 1174–1189.
- 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**, 1–12.
- Duran, B. S., Tsai, W. S. & Lewis, T. O. (1976), 'A class of location-scale non-parametric tests', *Biometrika* **63**, 173–176.
- Efron, B. (1979), 'Bootstrap methods: Another look at the jackknife', *Annals of Statistics* **7**, 1–26.

- Fisher, R. A. (1932), *Statistical Methods for Research Workers*, 4 edn, Oliver & Boyd, Edinburgh.
- Fleming, T. R., Harrington, D. P. & O'Sullivan, M. (1987), 'Supremum version of the logrank and generalized ilcoxon statistics', *Journal of the American Statistical Association* **82**, 312–320.
- Good, P. (2000), *Permutation Tests-A Practical Guide to Resampling Methods for Testing Hypotheses*, 2 edn, Springer, New York.
- Lepage, Y. (1971), 'A combination of Wilcoxon's and Ansari-Bradley's statistics', *Biometrika* **58**, 213–217.
- Lepage, Y. (1973), 'A table for a combined Wilcoxon Ansari-Bradley statistic', *Biometrika* **60**, 113–116.
- Liptak, I. (1958), 'On the combination of independent tests', *Magyar Tudományos Akademia Matematikai Kutato Intezenek Kozlomenyei* **3**, 127–141.
- Marozzi, M. (2004), 'A bi-aspect nonparametric test for the two-sample location problem', *Computational Statistics and Data Analysis* **44**, 639–648.
- Marozzi, M. (2007), 'Multivariate tri-aspect non-parametric testing', *Journal of Nonparametric Statistics* **19**, 269–282.
- Marozzi, M. (2009), 'Some notes on the location-scale Cucconi test', *Journal of Nonparametric Statistics* **21**, 629–647.
- Marozzi, M. (2012a), 'A modified Hall-Padmanabhan test for the homogeneity of scales', *Communications in Statistics-Theory and Methods* **41**, 3068–3078.
- Marozzi, M. (2012b), 'A combined test for differences in scale based on the interquantile range', *Statistical Papers* **53**, 61–72.
- Marozzi, M. (2012c), 'A distribution free test for the equality of scales', *Communications in Statistics-Simulation and Computation* **41**, 878–889.
- Marozzi, M. (2012d), 'A modified Cucconi test for location and scale change alternatives', *Revista Colombiana de Estadística* **35**, 371–384.
- Marozzi, M. (2013), 'Nonparametric simultaneous tests for location and scale testing: A comparison of several methods', *Communications in Statistics-Simulation and Computation* **42**, 1298–1317.
- Marozzi, M. (2014), 'The multisample Cucconi test', *Statistical Methods and Applications* **23**(2), 209–227.
- Mood, A. M. (1954), 'On the asymptotic efficiency of certain nonparametric two-sample tests', *Annals of Mathematical Statistics* **25**, 514–522.
- Murakami, H. (2007), 'Lepage type statistic based on the modified Baumgartner statistic', *Computational Statistics and Data Analysis* **51**, 5061–5067.

- Neuhäuser, 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**(3), 525–534.
- Park, H. I. (2011), 'A nonparametric test procedure based on a group of quantile tests', *Communications in Statistics-Simulation and Computation* **40**, 759–783.
- Pesarin, F. (2001), *Multivariate Permutation Tests*, Wiley, Chichester.
- Pesarin, F. & Salmaso, L. (2010), *Permutation Tests for Complex Data*, Wiley, Chichester.
- Pettitt, A. N. (1976), 'A two-sample Anderson-Darling rank statistic', *Biometrika* **63**, 161–168.
- Podgor, M. J. & Gastwirth, J. L. (1994), 'On non-parametric and generalized tests for the two-sample problem with location and scale change alternatives', *Statistics in Medicine* **13**, 747–758.
- Randles, R. H. & Hogg, R. V. (1971), 'Certain uncorrelated and independent rank statistics', *Journal of American Statistical Association* **66**, 569–574.
- Randles, R. H. & Wolfe, D. A. (1979), *Introduction to the Theory of Nonparametric Statistics*, Wiley, New York.
- Rublik, F. (2009), 'Critical values for testing location-scale hypothesis', *Measurement Science Review* **9**, 9–15.
- Salmaso, L. & Solari, A. (2005), 'Multiple aspect testing for case-control designs', *Metrika* **62**, 331–340.
- Shao, J. & Tu, D. (1995), *The Jackknife and Bootstrap*, Springer, New York.
- Tippett, L. H. C. (1931), *The Methods of Statistics*, Williams and Norgate, London.
- van Zwet, W. R. & Oosterhoff, J. (1967), 'On the combination of independent test statistics', *Annals of Mathematical Statistics* **38**, 659–680.
- Zhang, J. (2006), 'Powerful two-sample tests based on the likelihood ratio', *Technometrics* **48**, 95–103.