The multisample Cucconi test

Article *in* Statistical Methods and Applications · June 2014

1 author:

Marco Marozzi
Università Ca' Foscari Venezia
73 PUBLICATIONS **713** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

nonparametric charts View project

The multisample Cucconi test

Marco Marozzi

Abstract The multisample version of the Cucconi rank test for the two-sample location-scale problem is proposed. Even though little known, the Cucconi test is of interest for several reasons. The test is compared with some Lepage-type tests. It is shown that the multisample Cucconi test is slightly more powerful than the multisample Lepage test. Moreover, its test statistic can be computed analytically whereas several others cannot. A practical application example in experimental nutrition is presented. An R function to perform the multisample Cucconi test is given.

Keywords Nonparametric testing · The multisample location scale problem · The Cucconi test · Rank testing

1 Introduction

The usual two-sample testing problem is about location differences. In many biomedical situations, however, the treatment can change location and scale simultaneously [23]. The two-sample location-scale problem arises also in other fields like parasitology [28], human heredity [9], climate dynamics [1] and finance [14]. The standard rank test for the two-sample location-scale problem is the Lepage [12] test which is a combination of the Wilcoxon test for location and the Ansari-Bradley test for scale. Many rank tests have been proposed for this problem, quite all of them are Lepage-type tests, that is a combination of a location test and a scale test. [21], [6] and [27] proposed tests based on adaptive designs. [31], [26] and [24] proposed Lepage-type tests that are reviewed and compared by [20]. This comparison study also included a test by Cucconi [8] that is rather different than Lepage-type tests. The Cucconi test is little known because was published in Italian in a national journal, however it

Department of Environmental Sciences, Informatics and Statistics
University of Venice
E-mail: .

is of interest for several reasons. First, from a historical point of view, it is of interest because it was proposed before the Lepage test. Secondly, it is not a combination of a test for location and a test for scale. Thirdly, as shown by [16] and [20], it compares favorably with Lepage-type tests in terms of power, and its test statistic can be computed analytically whereas several others cannot due to challenging analytical computation.

After an oblivion of more than forty years, [16] put light on the Cucconi test computing for the very first time a table of exact critical values ([8] provided only asymptotic critical values) and studying its type-I error probability and power. The aim of this paper is to propose the multisample version of the Cucconi test. Section 2 reviews some tests for the multisample location-scale problem. The multisample Cucconi test is presented in Section 3. In Section 4, power estimation is addressed. Section 5 compares the multisample Cucconi test with the tests reviewed in Section 2. A practical application in experimental nutrition is discussed in Section 6. Concluding remarks are given in Section 7. Proofs of lemmas and theorems are reported in Appendix A. An R function to perform the multisample Cucconi test is given in Appendix B.

2 Review of some tests for the multisample location-scale problem

In this section, we review three rank tests for jointly detecting location and scale changes within the multisample setting. Suppose we have $K \geq 2$ independent random samples $\mathbf{X}_k = (X_{k1}, ..., X_{kn_k})$ from populations with continuous distribution functions $F_k(x) = F(\sigma_k x + \mu_k)$, $n_k \geq 2$, $k = 1, ..., K$. The function F, location parameters μ_k and scale parameters σ_k are unknown. We would like to address the multisample location-scale problem that is to test the null hypothesis

$$H_0 : (\mu_1, \sigma_1) = (\mu_2, \sigma_2) = ... = (\mu_K, \sigma_K)$$

against the alternative hypothesis

$$H_1 : H_0 \text{ not true.}$$

It is important to emphasize that the usual way to address the location-scale problem is the quadratic direct combination of a test for location and a test for scale. The corresponding test statistic is the sum of the squared standardized test statistics for location and for scale. When the null hypothesis is true, the location-scale test statistic is expected to be close to 0. When at least two of the distribution functions are different in location and/or in scale, the test statistic is larger than zero, the larger the difference in distributions, the larger the test statistic.

Let R_{ki} be the rank of X_{ki} in the pooled sample $\mathbf{X} = (\mathbf{X}_1, ..., \mathbf{X}_K)$, $n = n_1 + ... + n_K$. The multisample Lepage statistic [34] is the sum of the squared standardized multisample Wilcoxon [11] and Ansari-Bradley [33] statistics

$$ML = \frac{12}{n(n+1)} \sum_{k=1}^{K} n_k \left(W_k - \frac{n+1}{2} \right)^2 + MAB, \tag{1}$$

where

$$W_k = \frac{1}{n_k} \sum_{i=1}^{n_k} R_{ki},$$

$$MAB = \begin{cases} \frac{48(n-1)}{n(n^2-4)} \sum_{k=1}^{K} n_k \left(AB_k - \frac{n+2}{4}\right)^2 & \text{when } n \text{ is even} \\ \frac{48n^2}{n(n+1)(n^2+3)} \sum_{k=1}^{K} n_k \left(AB_k - \frac{(n+1)^2}{4n}\right)^2 & \text{when } n \text{ is odd} \end{cases}$$

and

$$AB_k = \frac{1}{n_k} \sum_{i=1}^{n_k} \left(\frac{n+1}{2} - \left|R_{ki} - \frac{n+1}{2}\right|\right).$$

We also consider the MH_1 and MH_2 tests by [25]. The MH_1 test is based on

$$MH_1 = \frac{[MB - E(MB)]^2}{Var(MB)} + \frac{180}{n(n+1)(n^2-4)} \sum_{k=1}^{K} n_k \left(M_k - \frac{n^2-1}{12}\right)^2, \quad (2)$$

where

$$MB = \frac{1}{K(n+1)} \sum_{k=1}^{K} \frac{(n_k+1)^2(n_k+2)}{n_k(n-n_k)} \sum_{i=1}^{n_k} \frac{\left(S_{ki} - \frac{n+1}{n_k+1}i\right)^2}{i(n_k+1-i)},$$

S_{ki} $i = 1, ..., n_k$ are the ranks of $\mathbf{X}_k$ in the pooled sample in increasing order, $E(MB)$ is the mean of MB under the null hypothesis, $Var(MB)$ is the variance of MB under the null hypothesis and

$$M_k = \frac{1}{n_k} \sum_{i=1}^{n_k} \left(R_{ki} - \frac{n+1}{2}\right)^2.$$

The MH_2 test is based on

$$MH_2 = \frac{[MB - E(MB)]^2}{Var(MB)} + MAB. \quad (3)$$

Note that both MH_1 and MH_2 are Lepage-type tests where the multisample Wilcoxon statistic is replaced with the multisample Baumgartner-Weiss-Schindler [3] statistic. In MH_1 the multisample Ansari-Bradley statistic is replaced with the multisample Mood [33] statistic. Note that large values of the ML, MH_1 and MH_2 statistics speak against the null hypothesis. The tests can be performed as permutation tests because under H_0 exchangeability of observations holds.

It is important to note that the multi-aspect framework proposed by [30] can be used for addressing the multisample location-scale problem. This framework can be used for location problems [15] and scale problems too [17], [18].

3 The multisample Cucconi test

We define the multisample Cucconi statistic as

$$MC = \frac{1}{K} \sum_{k=1}^{K} \frac{U_k^2 + V_k^2 - 2 U_k V_k \rho}{2 \left(1 - \rho^2 \right)}, \tag{4}$$

where

$$U_k = \frac{\sum_{i=1}^{n_k} R_{ki}^2 - E \left(\sum_{i=1}^{n_k} R_{ki}^2 \right)}{\sqrt{Var \left(\sum_{i=1}^{n_k} R_{ki}^2 \right)}}, \tag{5}$$

$$V_k = \frac{\sum_{i=1}^{n_k} \left(n + 1 - R_{ki} \right)^2 - E \left(\sum_{i=1}^{n_k} R_{ki}^2 \right)}{\sqrt{Var \left(\sum_{i=1}^{n_k} R_{ki}^2 \right)}}, \tag{6}$$

and

$$\rho = Cor(U_k, V_k) = Cor \left(\sum_{i=1}^{n_k} R_{ki}^2, \sum_{i=1}^{n_k} \left(n + 1 - R_{ki} \right)^2 \right). \tag{7}$$

$E(\sum_{i=1}^{n_k} R_{ki}^2)$ is given in Theorem 1, $Var \left(\sum_{i=1}^{n_k} R_{ki}^2 \right)$ in Theorem 2 and $Cor(U_k, V_k)$ in Theorem 3, see Appendix A.

Note that

$$E \left(\sum_{i=1}^{n_k} R_{ki}^2 \right) = E \left(\sum_{i=1}^{n_k} (n + 1 - R_{ki})^2 \right)$$

and that

$$Var \left(\sum_{i=1}^{n_k} R_{ki}^2 \right) = Var \left(\sum_{i=1}^{n_k} \left(n + 1 - R_{ki} \right)^2 \right).$$

Of course, $\sum_{i=1}^{n_k} R_{ki}^2$ and $\sum_{i=1}^{n_k} \left(n + 1 - R_{ki} \right)^2$ are highly negatively correlated, therefore to combine them we considered their correlation ρ.

The multisample Cucconi statistic is the natural extension of the corresponding two-sample statistic

$$C = \frac{U_1^2 + V_1^2 - 2 U_1 V_1 \rho}{2 \left(1 - \rho^2 \right)} = \frac{U_2^2 + V_2^2 - 2 U_2 V_2 \rho}{2 \left(1 - \rho^2 \right)}.$$

Note that if $K = 2$ then $MC = C$ because $U_1 = -U_2$ and $V_1 = -V_2$. For the two-sample case, [8] and [16] emphasized that under the null hypothesis that $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$ the bivariate random variable $(U_1, V_1) = (U_2, V_2)$ is centered on $(0, 0)$, whereas it is not under the alternative hypothesis. Therefore under H_0, observed values of the two-sample Cucconi C test statistic are expected to be close to 0, the larger the distance from 0, the stronger the evidence against H_0. More precisely, when $\mu_1 > \mu_2$ and $\sigma_1 = \sigma_2$, $\sum_{i=1}^{n_1} R_{1i}^2$ tends to be greater than $E(\sum_{i=1}^{n_1} R_{1i}^2)$ and so U_1 tends to be greater than 0 whereas V_1 tends to be less than 0. The contrary happens when $\mu_1 < \mu_2$. When $\mu_1 = \mu_2$ and $\sigma_1 > \sigma_2$, $\sum_{i=1}^{n_1} R_{1i}^2$ tends to be greater than $E(\sum_{i=1}^{n_1} R_{1i}^2)$

and then U_1 tends to be greater than 0 and so does V_1 because the ranks of the first sample elements tend to be the extreme elements of the sequence $1, 2, , n$, whereas when $\sigma_1 < \sigma_2$ they tend to be in the middle and then U_1 and V_1 tend to be less than 0. When $\mu_1 \neq \mu_2$ and $\sigma_1 \neq \sigma_2$, (U_1, V_1) is once again non-centered on $(0, 0)$, for example when $\mu_1 > \mu_2$ and $\sigma_1 < \sigma_2$, $E(U_1)$ may be close to 0 but $E(V_1)$ tends to be less than 0.

Note that the test statistic for the two-sample case can be seen as one half of the Mahalanobis distance between U_1 and V_1 or between U_2 and V_2

$$C = \frac{1}{2} \begin{bmatrix} U_1 \\ V_1 \end{bmatrix}^T \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}^{-1} \begin{bmatrix} U_1 \\ V_1 \end{bmatrix}.$$

Analogously, the multisample Cucconi MC test statistic can be seen as the mean of the Mahalanobis distance between U_k and V_k for $k = 1, ..., K$ divided by 2

$$MC = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{2} \begin{bmatrix} U_k \\ V_k \end{bmatrix}^T \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}^{-1} \begin{bmatrix} U_k \\ V_k \end{bmatrix}.$$

It is important to emphasize that the Mahalanobis distance takes appropriate account of the correlation ρ and is very different than the Euclidean distance because ρ is very high (in absolute value). From Theorem 3 in Appendix A, it is straightforward to show that ρ is an increasing function of n, $\rho \to -7/8$ as $n \to \infty$ and then $\rho < -7/8$ $\forall k$.

The null hypothesis is rejected (accepted) at the α nominal significance level if $MC \geq (<) mc_\alpha$, where the constant mc_α is chosen to make the type-I error probability equal to α. Alternatively one may perform the MC test as a permutation test, see the following algorithm.

1. Compute the MC statistic on the observed data set, denote it by MC_0;
2. pool together the samples, randomly permute the pooled sample and compute the MC statistic on the permuted data set, denote it by MC_1;
3. repeat steps 2 $B - 1$ times obtaining $MC_2, ..., MC_B$;
4. reject H_0 if the estimated p-value

$$\hat{p} = \frac{1}{B} \sum_{b=1}^{B} I(MC_b \geq MC_0)$$

is less than or equal to α and accept H_0 otherwise, where $I(.)$ denotes the indicator function.

See Appendix B for the corresponding R code. [19] combined the Lepage and Cucconi approaches and obtained a modified Cucconi test for the two-sample location-scale problem which compares favorably with the original one.

The possibility of performing the permutation test makes the derivation of the asymptotic distribution of the multisample Cucconi statistic not very important in practice. A reasonably accurate estimate of the exact p-value p of the test is obtained by considering a sufficiently large random sample of B permutations of the pooled sample. The larger B, the more accurate the

p-value, in general 1000 permutations are enough for testing at the $\alpha = .05$ level if the main purpose is the accept/reject decision on H_0. When the purpose is p-value estimation, we suggest to consider at least 10000 permutations, see the next section. Note that when all $C = \frac{n!}{\prod_{k=1}^{K} n_k!}$ possible permutations of the pooled sample are considered, then $\hat{p}$ would equal p. It is important to note that a reasonably accurate estimate of the exact p-value is enough to make inference on H_0 and often it is better than an asymptotic p-value since in practice samples sizes are generally small or moderate. Theoretically speaking, to find the asymptotic distribution is conversely interesting. For $K = 2$ [8] showed that under H_0 U_1 and V_1 are asymptotically normal and so are U_2 and V_2, and that $Cor(U_1, V_1) = Cor(U_2, V_2) = -7/8$. Therefore U_1^2, V_1^2, U_2^2, V_2^2 are asymptotically distributed as the chi square distribution with 1 degree of freedom. The C and MC statistics are asymptotically distributed as a quadratic form of dependent normal variables (Mahalanobis combination). Therefore, to find the asymptotic distribution of the MC statistic, one has to find the distribution of a combination of dependent chi square variables. This problem is very challenging since the easier problem of finding the distribution of a weighted sum of independent chi square variables has not yet a closed analytic expression [4]. Some possible solutions are: 1) fitting other distributions [13], 2) expanding the moment generating function in a series representation (e.g. Edgeworth or Laguerre series) and inverting term by term [7], 3) numerical integration to obtain the density function [2]. Unfortunately, these approaches are quite difficult and have some computational drawbacks too [4]: 1) there is no explicit expression of the error of fit, 2) prefactors in the expansion cannot be computed in polynomial time, 3) the number of dimensions over which we have to integrate for n terms is $n-1$ which is feasible only for small n.

4 Estimation of test power function

It is very difficult, if not even impossible, to derive theoretically optimality properties of a nonparametric test without any assumption on population distributions [10]. Therefore we estimate type-I error probability and power of the test using Monte Carlo simulations. Let SIM be the number of Monte Carlo simulations. Monte Carlo estimation of power function is obtained by generating SIM independent data sets and computing the proportion of rejections. The resulting root mean squared error $RMSE$ when the rejection probability 0.05 is estimated is $\sqrt{0.05(1-0.05)/SIM} = 0.00218$ when $SIM = 10000$. The maximum of $RMSE$ is when estimating a rejection probability of 0.5 and it is 0.005. In principle, by considering all C permutations we can compute the exact p-value p. In practice, unless the sample sizes are very small, it is computationally hard to consider all permutations, therefore we rely on a simple random sample of size B drawn with replacement from the C permutations and then compute the estimated p-value $\hat{p}$ as indicated in Section 3. The consequence is an inflated $RMSE$.

Table 1 Empirical power functions for tests on location-scale for $(n_1, n_2, n_3) = (5,5,5),(9,6,3)$, normal sampling and $\alpha = .05$.

μ_k		$(n_1,n_2,n_3) = (5,5,5)$ σ_k				$(n_1,n_2,n_3) = (9,6,3)$ σ_k			
		1 1 1	1 1 2	1 2 3	1 5 10	1 1 1	1 1 2	1 2 3	1 5 10
0 0 0	MC	0.049	0.144	0.181	0.437	0.046	0.167	0.357	0.849
	ML	0.048	0.130	0.178	0.483	0.046	0.146	0.285	0.761
	MH_1	0.050	0.136	0.162	0.424	0.045	0.175	0.338	0.829
	MH_2	0.048	0.121	0.160	0.481	0.047	0.129	0.278	0.782
0 1 1	MC	0.194	0.223	0.191	0.447	0.286	0.362	0.419	0.852
	ML	0.194	0.196	0.181	0.487	0.265	0.325	0.354	0.768
	MH_1	0.208	0.230	0.196	0.441	0.317	0.398	0.422	0.838
	MH_2	0.207	0.214	0.192	0.489	0.324	0.375	0.384	0.792
0 1 2	MC	0.518	0.418	0.245	0.447	0.564	0.502	0.467	0.861
	ML	0.503	0.398	0.230	0.475	0.563	0.478	0.412	0.773
	MH_1	0.556	0.402	0.262	0.438	0.597	0.525	0.486	0.846
	MH_2	0.552	0.392	0.256	0.481	0.603	0.510	0.453	0.804
0 2 3	MC	0.885	0.734	0.380	0.442	0.954	0.897	0.690	0.869
	ML	0.872	0.710	0.356	0.455	0.942	0.863	0.646	0.791
	MH_1	0.911	0.760	0.463	0.450	0.967	0.920	0.740	0.861
	MH_2	0.907	0.749	0.451	0.484	0.968	0.920	0.733	0.827

Power estimation concerns the accept/reject decision on H_0 that is different from p-value estimation. With this in mind, the aim is to find the optimal number of permutations to estimate the power function of the test given the Monte Carlo simulation size SIM. [29] showed that, given SIM, the optimal B, in the sense that it minimizes the maximum mean squared error (MSE) of the Monte Carlo power function estimator over α under the null hypothesis, is $B = 2\sqrt{SIM}$. Under the alternative, it is suggested $B = 4\sqrt{SIM}$. However, [5] suggest $B = 8\sqrt{SIM}$. In this case, when estimating a rejection probability of rej, the $RMSE$ is expected to be close to $1.2\sqrt{rej(1-rej)/SIM}$. Therefore $\max_{rej}(RMSE) = 0.6/\sqrt{SIM}$. Using 10000 Monte Carlo simulations, the maximum $RMSE$ when $B = 800$ is expected to be close to 0.006 and the $RMSE$ under the null hypothesis is expected to be close to 0.00262.

5 Comparison with other methods

In this section, we study type-I error probability and power of the multisample Cucconi test. As a term of reference we consider the multisample Lepage test, which is the natural extension of standard rank test for the location-scale problem. We also consider the other two Lepage-type tests reviewed in Section 2.

First of all it should be noted that the MC and ML statistics can be computed analytically whereas the MH_1 and MH_2 ones cannot and you have to perform a resampling based estimation of them. Some practitioners may prefer the MC and ML tests over the MH_1 and MH_2 tests for this reason even if nowadays the computational aspect is less important than in the past.

Table 2 Empirical power functions for tests on location-scale for (n_1, n_2, n_3) = $(20, 20, 20), (10, 20, 30)$, normal sampling and $\alpha = .05$.

| | | $(n_1, n_2, n_3) = (20, 20, 20)$ | | | | $(n_1, n_2, n_3) = (10, 20, 30)$ | | | |
| | | | σ_k | | | | σ_k | | |
μ_k		1 1 1	1 1 1.5	1 1.5 2	1 2 3	1 1 1	1 1 1.5	1 1.5 2	1 2 3
0 0 0	MC	0.050	0.252	0.485	0.888	0.049	0.267	0.290	0.631
	ML	0.049	0.204	0.407	0.821	0.053	0.194	0.265	0.627
	MH_1	0.048	0.244	0.485	0.896	0.048	0.235	0.286	0.673
	MH_2	0.048	0.207	0.423	0.848	0.050	0.210	0.275	0.654
0 0.5 0.5	MC	0.236	0.410	0.557	0.903	0.145	0.319	0.291	0.625
	ML	0.236	0.364	0.489	0.840	0.168	0.270	0.276	0.618
	MH_1	0.242	0.404	0.566	0.911	0.153	0.292	0.305	0.673
	MH_2	0.244	0.375	0.517	0.871	0.150	0.272	0.295	0.650
0 0.5 1	MC	0.650	0.675	0.693	0.920	0.569	0.599	0.439	0.665
	ML	0.649	0.640	0.638	0.871	0.562	0.545	0.405	0.639
	MH_1	0.662	0.651	0.693	0.926	0.583	0.558	0.450	0.702
	MH_2	0.661	0.631	0.659	0.894	0.579	0.541	0.437	0.677
0 1 1.5	MC	0.967	0.942	0.883	0.960	0.883	0.832	0.586	0.679
	ML	0.966	0.936	0.862	0.931	0.893	0.830	0.584	0.645
	MH_1	0.969	0.944	0.899	0.965	0.893	0.829	0.654	0.755
	MH_2	0.969	0.941	0.890	0.952	0.891	0.822	0.644	0.730

Table 3 Empirical power functions for tests on location-scale for (n_1, n_2, n_3) = $(5, 5, 5), (9, 6, 3)$, Cauchy sampling and $\alpha = .05$.

| | | $(n_1, n_2, n_3) = (5, 5, 5)$ | | | | $(n_1, n_2, n_3) = (9, 6, 3)$ | | | |
| | | | σ_k | | | | σ_k | | |
μ_k		1 1 1	1 1 2	1 2 3	1 5 10	1 1 1	1 1 2	1 2 3	1 5 10
0 0 0	MC	0.047	0.077	0.097	0.232	0.049	0.090	0.158	0.479
	ML	0.049	0.076	0.103	0.268	0.052	0.091	0.141	0.438
	MH_1	0.050	0.072	0.092	0.224	0.050	0.098	0.157	0.465
	MH_2	0.050	0.072	0.095	0.265	0.051	0.086	0.141	0.440
0 1 1	MC	0.095	0.110	0.107	0.243	0.139	0.161	0.202	0.492
	ML	0.096	0.109	0.106	0.278	0.130	0.153	0.185	0.447
	MH_1	0.116	0.122	0.110	0.231	0.153	0.184	0.206	0.478
	MH_2	0.115	0.123	0.112	0.270	0.157	0.175	0.194	0.454
0 1 2	MC	0.200	0.179	0.137	0.245	0.209	0.213	0.226	0.490
	ML	0.196	0.176	0.137	0.275	0.209	0.209	0.208	0.441
	MH_1	0.235	0.195	0.146	0.232	0.234	0.238	0.235	0.475
	MH_2	0.233	0.193	0.147	0.274	0.235	0.228	0.219	0.454
0 2 3	MC	0.349	0.300	0.184	0.245	0.435	0.409	0.331	0.522
	ML	0.350	0.302	0.182	0.275	0.417	0.389	0.304	0.477
	MH_1	0.414	0.342	0.225	0.248	0.479	0.453	0.359	0.515
	MH_2	0.416	0.342	0.226	0.283	0.484	0.454	0.352	0.496

Two more important aspects for comparing the tests are type-I error probability and power. Four groups of Monte Carlo experiments have been performed. Each group consisted of three samples with (n_1, n_2, n_3) =(5,5,5), (9,6,3), (20,20,20), (10,20,30). The data were generated from normal distributions with location μ_k and scale σ_k, Cauchy distributions with location μ_k and scale σ_k, uniform distributions in the interval $[0, a_k]$ and exponential distributions with rate λ_k, $k = 1, 2, 3$. $SIM = 10000$ values of each test statistic under each experimental situation were simulated and the proportions of re-

Table 4 Empirical power functions for tests on location-scale for $(n_1, n_2, n_3) = (20, 20, 20), (10, 20, 30)$, Cauchy sampling and $\alpha = .05$.

| | | $(n_1, n_2, n_3) = (20, 20, 20)$ | | | | $(n_1, n_2, n_3) = (10, 20, 30)$ | | | |
| | | | σ_k | | | | σ_k | | |
μ_k		1 1 1	1 1 1.5	1 1.5 2	1 2 3	1 1 1	1 1 1.5	1 1.5 2	1 2 3
0 0 0	MC	0.050	0.108	0.176	0.380	0.049	0.105	0.110	0.224
	ML	0.051	0.106	0.180	0.410	0.050	0.101	0.129	0.283
	MH_1	0.052	0.107	0.173	0.387	0.049	0.098	0.114	0.242
	MH_2	0.051	0.105	0.181	0.413	0.048	0.107	0.133	0.292
0 0.5 0.5	MC	0.106	0.158	0.200	0.412	0.070	0.124	0.119	0.221
	ML	0.106	0.157	0.203	0.431	0.077	0.123	0.132	0.269
	MH_1	0.111	0.163	0.206	0.421	0.077	0.124	0.130	0.247
	MH_2	0.112	0.161	0.211	0.437	0.078	0.131	0.143	0.284
0 0.5 1	MC	0.235	0.255	0.266	0.432	0.203	0.212	0.161	0.243
	ML	0.237	0.261	0.267	0.452	0.205	0.208	0.167	0.278
	MH_1	0.258	0.260	0.272	0.441	0.229	0.221	0.176	0.268
	MH_2	0.260	0.266	0.275	0.459	0.230	0.226	0.185	0.299
0 1 1.5	MC	0.480	0.451	0.387	0.492	0.348	0.309	0.210	0.261
	ML	0.488	0.458	0.383	0.498	0.383	0.328	0.215	0.285
	MH_1	0.526	0.480	0.419	0.515	0.404	0.349	0.259	0.308
	MH_2	0.532	0.484	0.422	0.525	0.410	0.360	0.260	0.323

Table 5 Empirical power functions for tests on location-scale for exponential sampling and $\alpha = .05$.

n_1	n_2	n_3					λ_k			
				1 1 1	1 1 5	1 3 5	1 5 5	1 5 10	1 10 10	1 10 20
5	5	5	MC	0.049	0.304	0.326	0.415	0.569	0.679	0.799
			ML	0.048	0.292	0.317	0.412	0.559	0.677	0.792
			MH_1	0.049	0.379	0.320	0.394	0.569	0.651	0.781
			MH_2	0.049	0.373	0.316	0.393	0.564	0.650	0.777
				1 1 1	1 1 5	1 3 5	1 5 5	1 5 10	1 10 10	1 10 20
9	6	3	MC	0.054	0.207	0.352	0.530	0.668	0.838	0.894
			ML	0.053	0.265	0.336	0.488	0.631	0.798	0.870
			MH_1	0.051	0.213	0.362	0.541	0.679	0.848	0.903
			MH_2	0.051	0.219	0.376	0.558	0.692	0.858	0.911
				1 1 1	1 1 1.5	1 1.5 2	1 2 2	1 2 3	1 3 3	1 2 4
20	20	20	MC	0.052	0.139	0.318	0.429	0.711	0.833	0.882
			ML	0.053	0.136	0.310	0.414	0.696	0.818	0.875
			MH_1	0.054	0.147	0.312	0.407	0.699	0.813	0.888
			MH_2	0.054	0.146	0.309	0.399	0.691	0.804	0.884
				1 1 1	1 1 1.5	1 1.5 2	1 2 2	1 2 3	1 3 3	1 2 4
10	20	30	MC	0.053	0.198	0.275	0.272	0.595	0.599	0.834
			ML	0.048	0.172	0.268	0.300	0.583	0.625	0.814
			MH_1	0.050	0.198	0.275	0.265	0.591	0.575	0.836
			MH_2	0.049	0.195	0.262	0.249	0.573	0.552	0.830

jections of the null hypothesis at the nominal .05 level were computed. The proportions that are obtained when the null hypothesis is true are estimates of type-I error probabilities of the tests. The proportions that are obtained when the null hypothesis is not true are estimates of power functions of the tests. The tests have been performed as permutation tests with $B = 1000$ which is practically as computational affordable as $B = 800$, see Section 4. $E(MB)$

Table 6 Empirical power functions for tests on location-scale for uniform sampling and $\alpha = .05$.

n_1	n_2	n_3					a_k			
				1 1 1	1 1 1.5	1 1.5 2	1 1.5 1.5	1 2 3	1 3 3	1 3 8
5	5	5	MC	0.055	0.151	0.192	0.105	0.350	0.299	0.773
			ML	0.054	0.150	0.186	0.107	0.335	0.293	0.756
			MH_1	0.054	0.135	0.192	0.112	0.396	0.428	0.812
			MH_2	0.056	0.133	0.189	0.112	0.387	0.425	0.804
				1 1 1	1 1 1.5	1 1.5 2	1 1.5 1.5	1 2 3	1 3 3	1 3 8
9	6	3	MC	0.051	0.160	0.318	0.216	0.566	0.691	0.874
			ML	0.049	0.159	0.300	0.193	0.548	0.671	0.866
			MH_1	0.050	0.158	0.344	0.231	0.625	0.755	0.907
			MH_2	0.049	0.139	0.331	0.229	0.622	0.766	0.909
				1 1 1	1 1 1.2	1 1.2 1.4	1 1.4 1.4	1 1.3 1.6	1 1.7 1.7	1 1.5 2
20	20	20	MC	0.049	0.207	0.378	0.417	0.618	0.781	0.886
			ML	0.052	0.178	0.345	0.408	0.584	0.794	0.879
			MH_1	0.050	0.192	0.397	0.493	0.661	0.874	0.929
			MH_2	0.050	0.173	0.374	0.483	0.642	0.874	0.924
				1 1 1	1 1 1.2	1 1.2 1.4	1 1.4 1.4	1 1.3 1.6	1 1.7 1.7	1 1.5 2
10	20	30	MC	0.048	0.216	0.254	0.134	0.437	0.242	0.717
			ML	0.049	0.171	0.226	0.169	0.414	0.336	0.706
			MH_1	0.049	0.194	0.247	0.172	0.451	0.388	0.775
			MH_2	0.049	0.186	0.240	0.179	0.446	0.414	0.771

Table 7 Empirical power functions for tests on location-scale for $(n_1, n_2, n_3) = (20, 20, 20)$, binomial sampling and $\alpha = .05$.

		$(n_1, n_2, n_3) = (20, 20, 20)$				$(n_1, n_2, n_3) = (10, 20, 30)$			
			σ_k				σ_k		
μ_k		1 1 1	1 1 1.5	1 1.5 2	1 2 3	1 1 1	1 1 1.5	1 1.5 2	1 2 3
0 0 0	MC	0.049	0.294	0.522	0.787	0.050	0.290	0.334	0.515
	ML	0.050	0.337	0.548	0.735	0.053	0.338	0.420	0.561
	MH_1	0.049	0.339	0.618	0.874	0.053	0.334	0.429	0.637
	MH_2	0.049	0.375	0.631	0.844	0.054	0.400	0.493	0.653
0 0.5 0.5	MC	0.100	0.341	0.595	0.903	0.080	0.309	0.346	0.639
	ML	0.102	0.356	0.576	0.893	0.089	0.356	0.425	0.746
	MH_1	0.138	0.438	0.699	0.947	0.103	0.389	0.471	0.775
	MH_2	0.142	0.448	0.686	0.944	0.106	0.438	0.519	0.814
0 0.5 1	MC	0.332	0.624	0.691	0.912	0.279	0.569	0.418	0.644
	ML	0.339	0.557	0.654	0.802	0.283	0.497	0.384	0.598
	MH_1	0.453	0.706	0.798	0.959	0.386	0.657	0.593	0.798
	MH_2	0.454	0.670	0.780	0.917	0.389	0.633	0.571	0.751
	MH_2	0.454	0.670	0.780	0.917	0.389	0.633	0.571	0.751
0 1 1.5	MC	0.854	0.915	0.893	0.952	0.682	0.797	0.587	0.669
	ML	0.855	0.893	0.881	0.881	0.707	0.771	0.575	0.577
	MH_1	0.910	0.953	0.954	0.982	0.777	0.874	0.799	0.854
	MH_2	0.910	0.946	0.950	0.967	0.778	0.861	0.789	0.803

and $Var(MB)$ are estimated using a random sample of 1000 permutations of the pooled sample. Tables 1 to 6 display the results.

All the tests have a type-I error probability quite close to the nominal significance level because the minimum estimated significance levels are .046, .046, .045, .046 respectively for the MC, ML, MH_1, MH_2 tests and the maximum estimated significance levels are .055, .054, .054, .056 respectively.

In terms of power, the MC, MH_1 and MH_2 tests perform very similarly and slightly better than the ML test when sampling from the normal distribution. The MH_1 and MH_2 tests perform similarly when sampling from the Cauchy distribution and very similarly when sampling from the uniform distribution. In both cases, the MC and ML tests perform very similarly and slightly worse than the MH_1 and MH_2 tests. When sampling from the exponential distribution, all the tests perform similarly. In summary, the tests perform quite similarly in terms of power, with the MH_1 and MH_2 tests that tend to have slightly higher power than the MC test and especially the ML test.

Albeit we assumed that distributions are continuous, in practice data modification, rounding and truncating can lead to ties. We address ties by assigning average ranks. A second round of simulations with ties has been performed. Tied data have been simulated using a binomial distribution with $(15, 0.5)$ parameters, $(n_1, n_2, n_3) = (20, 20, 20), (10, 20, 30)$, relocated and rescaled so that the mean is 0 and the standard deviation is 1. Table 7 displays the results. The type-I error probabilities of the tests range from 0.049 to 0.054. The MH_1 and MH_2 tests perform similarly and are more powerful than both the MC and ML tests.

6 Practical application

Food and nutrition play a central role in public health. Strong links between food and health have been shown by epidemiological evidence. It is well known that the risk of developing cardiovascular diseases (which are the first cause of death) depends on blood pressure levels. Many agencies, like the World Health Organization, have acknowledged the role of lifestyle and diet on blood pressure levels. In particular, high sodium intake is linked to a number of health problems because it causes hypertension, a primary risk factor for cardiovascular and renal diseases. Several studies estimated the potential societal savings and quality of life improvements of a reduced sodium intake. For example, [32] estimated that reducing average US population sodium intake to the recommended maximum for adults (that is 2300 mg per day according to the Centers for Disease Control and Prevention, US Department of Health and Human Services) may reduce cases of hypertension by 11 million and save 18 billion dollars in health care expenses. Sodium reduction is one of the most cost effective and easily implemented strategies to improve population health. In the past, the diet of humans consisted of plant and animal foods containing small amounts of sodium. Today, many people consume far more sodium because especially in the developed countries the diet has a large share of processed foods which account for the most majority of dietary salt. One of the most popular processed food in the world is the hot dog. Table 8 displays the milligrams of sodium per hot dog in each of 54 hot dog brands in the USA. Hot dogs are classified by type: beef, poultry and meat (mostly pork and beef). The data source is [22]. We are interested in assessing whether there are significant

Table 8 Results of a laboratory analysis of sodium content (mg) of 54 hot dog brands.

				Beef					
495	477	425	322	482	587	370	322	479	375
330	300	386	401	645	440	317	319	298	253
				Meat					
458	506	473	545	496	360	387	386	507	393
405	372	144	511	405	428	339	-	-	-
				Poultry					
430	375	396	383	387	542	359	357	528	513
426	513	358	581	588	522	545	-	-	-

differences in location and scale between the types of hot dog. We considered a random sample of 1000000 permutations of the pooled sample and we computed the p-values of the tests studied before. The results are 0.075 (MC), 0.081 (ML), 0.102 (MH_1) and 0.104 (MH_2). Therefore all the tests do not find enough evidence in the data for rejecting at the 0.05 nominal significance level the null hypothesis that locations and scales of sodium content are the same irrespective of hot dog type. R code for analyzing hot dog data using the MC test is reported in Appendix B.

7 Conclusion

The Cucconi test for the two-sample location-scale problem is little known. It is of interest from a historical point of view because has been proposed before the Lepage test, which is the standard rank test for the location-scale problem. It is also of interest because is not a combination of a test for location and a test for scale as the other location-scale tests, and because, as shown by [16] and [20], it compares favorably with Lepage-type tests in terms of power. We proposed the multisample version of the Cucconi test and compared it with some Lepage-type tests: the multisample Lepage test and two tests proposed by [25]. It is shown that the tests perform quite similarly in terms of power. An R function for performing the multisample Cucconi test as a permutation test is reported in Appendix B. We suggest to address the multisample location-scale problem through the Cucconi test which is slightly more powerful than the Lepage test. Slightly higher power (especially in the presence of many ties) can be obtained at the expense of more complicated procedures using Lepage-type tests proposed by [25].

8 Appendix A

[8] does not organize the formal results on $\sum_{i=1}^{n_k} R_{ki}^2$ and $\sum_{i=1}^{n_k} \left(n + 1 - R_{ki} \right)^2$ in lemmas and theorems but gives an outline of some derivations. Here we reorganize the results on the sums of squared ranks and squared antiranks in a clearer manner by reporting the whole proof of all results. Well known

results about sum of powers of the first n natural numbers will be used in this section.

Theorem 1 *Under the null hypothesis*

$$E\left(\sum_{i=1}^{n_k} R_{ki}^2\right) = n_k (n+1)(2n+1)/6 \ \forall k.$$

Proof Consider the population of the first n squared natural numbers $1, 2^2, ..., n^2$. $\frac{1}{n_k}\sum_{i=1}^{n_k} R_{ki}^2$ may be seen as the random variable defined by the mean of a random sample of n_k values drawn without replacement from this population. Since the mean of the sample means is equal to the population mean it follows that

$$E\left(\frac{1}{n_k}\sum_{i=1}^{n_k} R_{ki}^2\right) = \frac{1}{n}\sum_{i=1}^{n} i^2 = (n+1)(2n+1)/6$$

$\forall k$, and the thesis follows immediately.

Theorem 2 *Under the null hypothesis*

$$Var\left(\sum_{i=1}^{n_k} R_{ki}^2\right) = n_k (n-n_k)(n+1)(2n+1)(8n+11)/180 \ \forall k.$$

Proof With simple algebra we first compute the variance τ^2 of the population of the first n squared natural numbers

$$\tau^2 = \frac{1}{n}\sum_{i=1}^{n} \left[i^2 - (2n+1)(n+1)/6\right]^2$$

$$= (n^2-1)(2n+1)(8n+11)/180.$$

Now, since the sampling is without replacement

$$Var\left(\frac{1}{n_k}\sum_{i=1}^{n_k} R_{ki}^2\right) = \frac{n-n_k}{n-1}\frac{\tau^2}{n_k} \tag{8}$$

$\forall k$, and the thesis follows immediately.

The following lemmas will be used for proving Theorem 3.

Lemma 1 *Under the null hypothesis*

$$E\left[\left(\sum_{i=1}^{n_k} R_{ki}^2\right)^2\right] = n_k (n-n_k)(n+1)(2n+1)(8n+11)/180 +$$

$$+n_k^2 (n+1)^2 (2n+1)^2 /36 \ \forall k.$$

Proof Straightforward (we deliberately not factor the result).

Lemma 2 *Under the null hypothesis*

$$E\left(\sum_{i=1}^{n_k} R_{ki}^2 \sum_{i=1}^{n_k} R_{ki}\right) = n_k n\left(n+1\right)^2\left(2n_k+1\right)/12 \ \forall k.$$

Proof

$$E\left(\sum_{i=1}^{n_k} R_{ki}^2 \sum_{i=1}^{n_k} R_{ki}\right) = E\left(\sum_{i=1}^{n_k} R_{ki}^3\right) + E\left(\sum_{i=1}^{n_k}\sum_{j\neq 1}^{n_k} R_{ki}^2 R_{kj}\right).$$

It is $E\left(\sum_{i=1}^{n_k} R_{ki}^3\right) = \frac{n_k}{n}\sum_{j=1}^{n} j^3 = n_k n\left(n+1\right)^2/4$ and it is

$$\begin{aligned}
E\left(\sum_{i=1}^{n_k}\sum_{j\neq 1}^{n_k} R_{ki}^2 R_{kj}\right) &= \frac{n_k(n_k-1)}{n\left(n-1\right)}\sum_{j=1}^{n} j^2 \sum_{l\neq j}^{n} l \\
&= \frac{n_k(n_k-1)}{n\left(n-1\right)}\left(\sum_{j=1}^{n} j^2 \sum_{l=1}^{n} l - \sum_{j=1}^{n} j^3\right) \\
&= n_k(n_k-1)n\left(n+1\right)^2/6.
\end{aligned}$$

Finally it follows with simple algebra that

$$E\left(\sum_{i=1}^{n_k} R_{ki}^2 \sum_{i=1}^{n_k} R_{ki}\right) = n_k n\left(n+1\right)^2\left(2n_k+1\right)/12$$

$\forall k$.

Theorem 3 *Under the null hypothesis*

$$Cor\left(\sum_{i=1}^{n_k} R_{ki}^2, \sum_{i=1}^{n_k}\left(n+1-R_{ki}\right)^2\right) = -\frac{30n+14n^2+19}{\left(8n+11\right)\left(2n+1\right)} \ \forall k.$$

Proof It is

$$\begin{aligned}
&Cor\left(\sum_{i=1}^{n_k} R_{ki}^2, \sum_{i=1}^{n_k}\left(n+1-R_{ki}\right)^2\right) \\
&= \frac{E\left[\left(\sum_{i=1}^{n_k} R_{ki}^2\right)\left(\sum_{i=1}^{n_k}\left(n+1-R_{ki}\right)^2\right)\right] - n_k^2\left(2n+1\right)^2\left(n+1\right)^2/36}{n_k\left(n-n_k\right)\left(n+1\right)\left(2n+1\right)\left(8n+11\right)/180}.
\end{aligned}$$

It remains to compute

$$E\left[\left(\sum_{i=1}^{n_k} R_{ki}^2\right)\left(\sum_{i=1}^{n_k}(n+1-R_{ki})^2\right)\right]$$

$$= n_k(n+1)^2 E\left(\sum_{i=1}^{n_k} R_{ki}^2\right) + E\left[\left(\sum_{i=1}^{n_k} R_{ki}^2\right)^2\right] - 2(n+1)E\left(\sum_{i=1}^{n_k} R_{ki}^2\sum_{i=1}^{n_k} R_{ki}\right).$$

Using Lemma 1 it follows that

$$Cov\left(\sum_{i=1}^{n_k} R_{ki}^2, \sum_{i=1}^{n_k}(n+1-R_{ki})^2\right)$$

$$= n_k(n+1)^2 E\left(\sum_{i=1}^{n_k} R_{ki}^2\right) + Var\left(\sum_{i=1}^{n_k} R_{ki}^2\right) - 2(n+1)E\left(\sum_{i=1}^{n_k} R_{ki}^2\sum_{i=1}^{n_k} R_{ki}\right).$$

Using Theorem 1, Theorem 2 and Lemma 2 it follows with simple algebra that

$$Cov\left(\sum_{i=1}^{n_k} R_{ki}^2, \sum_{i=1}^{n_k}(n+1-R_{ki})^2\right)$$

$$= -n_k(n+1)\left(30n+14n^2+19\right)(n-n_k)/180$$

and finally that

$$Cor\left(\sum_{i=1}^{n_k} R_{ki}^2, \sum_{i=1}^{n_k}(n+1-R_{ki})^2\right)$$

$$= -\frac{30n+14n^2+19}{(8n+11)(2n+1)}$$

$\forall k$.

9 Appendix B

An R function for performing the multisample Cucconi test follows.

```
MultiSampleCucconiTest <- function(pooled.sample, sample.sizes, B){
  # Outer function: perform a permutation test using test statistic values
  #                 computed using the inner function
  #
  # Args:
  #   pooled.sample: the pooled sample
  #   sample.sizes: the vector with sample sizes
  #   B: the number of permutations
  #
  # Returns:
  #   mc.test[1]: the observed MC test statistic
  #   mc.test[2]: the estimated p-value of the MC test
  MultiSampleCucconiStat <- function(pooled.sample, sample.sizes){
```

```
  # Inner function: computes the MC test statistic
  #
  # Args:
  #   pooled.sample: the pooled sample
  #   sample.sizes: the vector with sample sizes
  #
  # Returns:
  #   The MC test statistic
  sample.sizes <- c(0, sample.sizes)
  n <- sum(sample.sizes)
  K <- length(sample.sizes)
  ranks <- rank(pooled.sample)
  contrary.ranks <- n+1-ranks
  groups.ranks <- vector("list", K-1)
  groups.contrary.ranks <- vector("list", K-1)
  cum.sample.sizes <- cumsum(sample.sizes)

  for (k in 1:(K-1)){
    groups.ranks[[k]] <- ranks[(cum.sample.sizes[k]+1):(cum.sample.sizes[k+1])]
    groups.contrary.ranks[[k]] <- contrary.ranks[(cum.sample.sizes[k]+1):(cum.sample.sizes[k+1])]
  }

  sample.sizes <- sample.sizes[2:length(sample.sizes)]
  means <- vector(, K-1)
  st.dev <- vector(, K-1)

  for (k in 1:(K-1)){
    means[k] <- sample.sizes[k]*(n+1)*(2*n+1)/6
    st.dev[k] <- (sample.sizes[k]*(n-sample.sizes[k])*(n+1)*(2*n+1)*(8*n+11)/180)^0.5
  }

  covariance <- -(30*n+14*n^2+19)/(8*n+11)/(2*n+1)

  u.stat <- vector(, K-1)
  v.stat <- vector(, K-1)
  partial.c.stat <- vector(, K-1)

  for (k in 1:(K-1)){
    u.stat[k] <- (sum(as.vector(groups.ranks[[k]])^2)-means[k])/st.dev[k]
    v.stat[k] <- (sum(as.vector(groups.contrary.ranks[[k]])^2)-means[k])/st.dev[k]
    partial.c.stat[k] <- (u.stat[k]^2+v.stat[k]^2-2*u.stat[k]*v.stat[k]*covariance)/2/(1-covariance^2)
  }

  c.stat <- mean(partial.c.stat)
  return(c.stat)
}  # end of the inner function

mc.test <- vector(, 2)
permutation.cucconi.stat <- vector(, B)

for (b in 1:B){
  pooled.sample.perm <- sample(pooled.sample)
  permutation.cucconi.stat[b] <- MultiSampleCucconiStat(pooled.sample.perm, sample.sizes)
}

mc.test[1] <- MultiSampleCucconiStat(pooled.sample, sample.sizes)
mc.test[2] <- length(permutation.cucconi.stat[permutation.cucconi.stat >= mc.test[1]])/B
return(mc.test)
```

```
}
```

To analyze the data considered in Section 6 run the following code.

```
# beef
x1=c(495,477,425,322,482,587,370,322,479,375,330,300,386,401,645,440,317,319,298,253)
# poultry
x2=c(458,506,473,545,496,360,387,386,507,393,405,372,144,511,405,428,339)
# meat
x3=c(430,375,396,383,387,542,359,357,528,513,426,513,358,581,588,522,545)

pooled.sample <- c(x1, x2, x3)
sample.sizes <- c(length(x1), length(x2), length(x3))
MultiSampleCucconiTest(pooled.sample, sample.sizes, 1000000)
```

References

1. Adamson GCD, Nash DJ (2013) Long-term variability in the date of monsoon onset over western India, Climate Dynamics, 40, 2589–2603.
2. Akkouchi M (2005) On the convolution of gamma distributions, Soochow Journal of Mathematics, 31, 205–211.
3. Baumgartner W, Weiss P, Schindler H (1998) A nonparametric test for the general two-sample problem, Biometrics, 54, 1129–1135.
4. Bausch J (2012) On the efficient calculation of a linear combination of chi-square random variables with an application in counting string vacua, arXiv:1208.2691v2.
5. Boos DD, Zhang J (2000) Monte Carlo evaluation of resampling-based hypothesis tests, Journal of the American Statistical Association, 95, 486–492.
6. Buning H, Thadewald T (2000) An adaptive two-sample location-scale test of Lepage-type for Symmetric distributions, Journal of Statistical Computation and Simulation, 65, 287–310.
7. Castano-Martinez A, Lopez-Blazquez F (2005) Distribution of a sum of weighted non-central chi-square variables, Test, 14, 397–415.
8. Cucconi O (1968) Un nuovo test non parametrico per il confronto tra due gruppi campionari, Giornale degli Economisti, 27, 225–248.
9. Gerhard D, Hothorn LA (2010) Rank transformation in haseman-elston regression using scores for location-scale alternatives, Human Heredity, 69, 143–151.
10. Hajek J, Sidak Z, Sen PK (1998). Theory of Rank Tests 2nd edition. New York: Academic Press.
11. Kruskal WH, Wallis WA (1952) Use of ranks in one criterion analysis of variance, Journal of the American Statistical Association, 47, 583–621.
12. Lepage Y (1971) A combination of Wilcoxon's and Ansari-Bradley's statistics, Biometrika, 58, 213–217.
13. Lindsay BG, Pilla RS, Basak P (2000) Moment-based approximations of distributions using mixtures: theory and applications, Annals of the Institute of Mathematical Statistics, 52, 215–230.
14. Lunde A, Timmermann A (2004) Duration dependence in stock prices: an analysis of bull and bear markets, Journal of Business and Economic Statistics, 22, 253–273.
15. Marozzi M (2007) Multivariate tri-aspect non-parametric testing, Journal of Nonparametric Statistics, 19, 269–282.
16. Marozzi M (2009) Some notes on the location-scale Cucconi test, Journal of Nonparametric Statistics, 21, 629–647.
17. Marozzi M (2012a) A combined test for differences in scale based on the interquantile range, Statistical Papers, 53, 61–72.

18. Marozzi M (2012b) A modified Hall-Padmanabhan test for the homogeneity of scales, Communications in Statistics - Theory and Methods, 41, 16–17, 3068–3078.
19. Marozzi M (2012c) A modified Cucconi test for location and scale change alternatives, Colombian Journal of Statistics, 35, 369–382.
20. Marozzi M (2013) Nonparametric simultaneous tests for location and scale testing: a comparison of several methods, Communication in Statistics - Simulation and Computation, 42, 6, 1298–1317.
21. Manly BFJ, Francis RICC (2002) Testing for mean and variance differences with samples from distributions that may be non-normal with unequal variances, Journal of Statistical Computation and Simulation, 72, 8, 633–646.
22. Moore DS, McCabe GP (2009) Introduction to the Practice of Statistics, 6th edition. New York: Freeman.
23. Muccioli C, Belford R, Podgor M, Sampaio P, de Smet M, Nussenblatt R (1993) The diagnosis of intraocular inflammation and cytomegalovirus retinitis in HIV-infected patients by laser flare photometry, Ocular Immunology and Inflammation, 4, 2, 75–81.
24. Murakami H (2007) Lepage-type statistic based on the modified Baumgartner statistic, Computational Statistics and Data Analysis, 51, 5061–5067.
25. Murakami H (2008) A multisample rank test for location-scale parameters, Communication in Statistics - Simulation and Computation, 37, 1347–1355.
26. Neuhauser M (2000) An exact two-sample test based on the Baumgartner-Weiss-Schindler statistic and a modification of Lepages test, Communication in Statistics Theory and Methods, 29, 67–78.
27. Neuhauser M (2001) An adaptive location-scale test, Biometrical Journal, 43, 809–819.
28. Neuhauser M, Kotzmann J, Walier M, Poulin R (2010) The comparison of mean crowding between two groups, Journal of Parasitology, 96, 477–481.
29. Oden NL (1991) Allocation of effort in Monte Carlo simulation for power of permutation tests, Journal of the American Statistical Association, 86, 1074–1076.
30. Pesarin F, Salmaso L (2010) Permutation Tests for Complex Data. Chichester: Wiley.
31. Podgor MJ, Gastwirth JL (1994) On non-parametric and generalized tests for the two-sample problem with location and scale change alternatives, Statistics in Medicine, 13, 747–758.
32. Palar K, Sturm R (2009) Potential societal savings from reduced sodium consumption in the U.S. adult population, Americal Journal of Health Promotion, 24, 49–57.
33. Puri ML (1965) On some tests of homogeneity of variances, Annals of the Institute of Statistical Mathematics, 17, 323–30.
34. Rublik F (2005) The multisample version of the Lepage test, Kybernetika, 41, 713–733.