***S**tatistica & **A**pplicazioni*
Vol. IV, special issue n° 1, 2006
pp. 33-47

Some developments about a new nonparametric test based on Gini's mean difference

Claudio Giovanni Borroni §

Manuela Cazzaro §

Summary: *In this paper the performance of a new nonparametric test proposed by Borroni and Zenga (2003) for the independence of two criteria is discussed. The test-statistic is based on Gini's mean difference computed on the total ranks assigned to each sampled unit according to the chosen criteria of sorting. The performance of the test is measured by simulating its power function via Monte Carlo methods when it is applied as a one-sided test of independence against concordance. At this aim, after assuming that the two criteria of ranking are based on the values taken by two quantitative variables, a wide range of bivariate models is set for the two populations. The choice of the simulated models reflects the common situation of sampling from non-Normal populations, usually faced in economic applications. The reported results show that the new proposed test has often good performances and can be considered as a natural competitor of other common nonparametric tests.*

Keywords: *nonparametric tests, rank correlation indexes, Gini's mean difference, Monte Carlo simulations.*

1. Introduction

Consider a sample of *n* units which are ranked according to two criteria or to the values taken by two quantitative variables. For instance, *n* students can be ranked according to their ability in two different subjects, mathematics and music. A common statistical problem consists of testing the hypothesis of independence of the two criteria of ranking; the rejection of such hypothesis will lead to conclude, for instance, that the abilities of students in

§ Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali – Università degli Studi di Milano-Bicocca – piazza dell'Ateneo Nuovo, 1, 20126 MILANO (e-mail: claudio.borroni@unimib.it, manuela.cazzaro@unimib.it).

the two subjects are related or that the two variables used for ranking the sampled units are dependent. Moreover, the researcher will often want to test independence against a specific alternative - concordance or discordance of the two criteria - rather than against the absence of independence. For instance, it may be interesting to test if students with high ability in maths tend to show a bad attitude to music. When the criteria of ranking are based on two quantitative variables X and Y with joint cdf $H(x, y)$, the concepts of independence, concordance and discordance can be more analytically specified. In particular, the null hypothesis of the test can be stated as H_0: $H(x, y) = F(x)\cdot G(y)$ (for all $(x, y)\in\Re^2$) where $F(x)$ and $G(y)$ denote the marginal cdf's of X and Y respectively. Concerning the alternative hypothesis, a specific bivariate dependence concept has to be chosen. In these contexts, it is common to refer to the *positive/negative quadrant dependence*, i.e. H_0 can be tested against the alternative of concordance H_1^+: $H(x, y) \geq F(x)\cdot G(y)$ (with strict inequality for at least one couple (x, y)) or against the other one-sided alternative hypothesis H_1^-: $H(x, y) \leq F(x)\cdot G(y)$ (with strict inequality for at least one couple (x, y)), which represents discordance of the criteria. Moreover, the two-sided alternative $\mathrm{H}_1 = \mathrm{H}_1^+\cup\mathrm{H}_1^-$ can be selected.

After ranking the n sampled units according to the selected criteria, the two corresponding sequences of ranks will be denoted by $R_{11}, R_{21}, \ldots, R_{n1}$ and $R_{12}, R_{22}, \ldots, R_{n2}$. Notice that, in the sample, two opposite extreme situations corresponding to the complete agreement or disagreement of the two criteria of ranking can be observed. When the two criteria are perfectly concordant, $R_{i1} = R_{i2}$ $(i = 1, \ldots, n)$, whereas $R_{i1} = n+1-R_{i2}$ $(i = 1, \ldots, n)$, in case of perfect discordance. Of course any observed situation will always lie between these two extremes. Any index measuring how the observed situation is far from its extremes may be termed as a *rank correlation index* (see Kendall and Gibbons (1990)). In the literature of such indexes, the possibility of ties in the sample is usually ignored; this fact can be justified, when the sampled units are ranked according to two quantitative variables, by assuming their continuity. In a more general context, the occurrence of ties is often treated by using mid-ranks; even if this problem can be an interesting issue, this work will not deal with it.

The most famous rank correlation indexes are Spearman's rho

$$\rho = \frac{12}{n^3 - n}\sum_{i=1}^{n}\left(R_{i1} - \frac{n+1}{2}\right)\left(R_{i2} - \frac{n+1}{2}\right) = 1 - \frac{6\sum_{i=1}^{n}\left(R_{i1} - R_{i2}\right)^2}{n^3 - n} \qquad (1)$$

and Kendall's tau

$$\tau = 1 - \frac{4Q}{n^2 - n} \qquad (2)$$

where Q denotes the number of discordant pairs of ranks, i.e. the number of pairs (R_{i1}, R_{i2}) and (R_{l1}, R_{l2}) such that $R_{i1} < R_{i2}$ and $R_{l1} > R_{l2}$ or $R_{i1} > R_{i2}$ and $R_{l1} < R_{l2}$ ($i,l = 1, \ldots, n$). Another measure of rank correlation is Gini's cograduation index:

$$G = \frac{2}{g} \sum_{i=1}^{n} \left\{ \left| R_{i1} + R_{i2} - n - 1 \right| - \left| R_{i1} - R_{i2} \right| \right\} \qquad (3)$$

(Gini (1954); see also Cifarelli et al. (1996)) where the normalization constant g equals n^2 when n is even and n^2-1 when n is odd.

Recently, a new rank correlation index has been introduced by Borroni and Zenga (2003). Denote by $T_i = R_{i1} + R_{i2}$ ($i = 1, \ldots, n$) the total ranks corresponding to each sampled unit. Now note that, when the two sequences of ranks are perfectly discordant, $T_i = n+1$ ($i = 1, \ldots, n$), that is the sequence of totals is constant. Conversely, the more the observed situation departs from discordance, the more the totals $T_1, \ldots, T_n$ show high variability. In the opposite extreme case of perfect concordance $T_i = 2 \cdot R_{i1} = 2 \cdot R_{i2}$ ($i = 1, \ldots, n$), that is the sequence of totals corresponds to $\{2, 4, 6, \ldots, 2n\}$ (in a convenient order).[1]

Hence if a suitable index of variability is computed on the totals $T_1, \ldots, T_n$, it will measure the concordance of the two sequences of ranks $R_{11}, R_{21}, \ldots, R_{n1}$ and $R_{12}, R_{22}, \ldots, R_{n2}$. Borroni and Zenga (2003) propose to use *Gini's mean difference* as a measure of variability (Gini (1912); see also David (1968) and Kotz and Johnson (1982), p. 436-437).

For N observations $x_1, \ldots, x_N$, relating to a quantitative variable X, Gini's mean difference (without repetition) can be defined as:

$$\Delta(X) = \frac{1}{N^2 - N} \sum_{i=1}^{N} \sum_{l=1}^{N} \left| x_i - x_l \right| \qquad (4)$$

that is as the mean of the $N(N-1)$ absolute differences between every couple of different observations. If this measure of variability is applied to the totals $T_1, \ldots, T_n$, the following normalized rank correlation index is obtained:

[1] This conclusions are highly related to the wider concept of *compensation* developed by Zenga (2003) (see the references for details).

$$D = \frac{3\Delta(T)}{n+1} - 1 \qquad (5)$$

where

$$\Delta(T) = \frac{1}{n^2 - n} \sum_{i=1}^{n} \sum_{l=1}^{n} |T_i - T_l|.$$

Note that, in case of perfect discordance, $\Delta(T) = 0$ and hence $D = -1$; conversely, in case of perfect concordance $\Delta(T) = (2/3)(n+1)$ and hence $D = +1$. As a result, D ranges in $[-1, 1]$ like the other rank correlation indexes above.

In Borroni and Zenga (2003) the analytical properties of D are examined. In particular, it is shown that, as a sample measure, D agrees with the so called *quadrant dependence ordering* according to which the bivariate cdf $H_2(x, y)$ is more concordant than $H_1(x, y)$ whenever $H_1(x, y) \leq H_2(x, y)$ for every couple $(x, y) \in \Re^2$. Hence, as below detailed, along with ρ, τ, and G, the index D can be conveniently used as a test-statistic for testing H_0 against one of the alternatives H_1, H_1^+ or H_1^-.

In this framework, it is natural to ask which of a proposed set of rank correlation indexes performs better as a testing procedure. This work will then aim at simulating and comparing the power functions of the tests based on D with the one corresponding to ρ, τ and G. This comparison is obviously far from giving a complete view of the properties of D as a test-statistic of independence. However, we feel that this study can give some useful guidelines for a correct use of D, especially in economic applications. Regarding the general characteristics of D, the reader may instead refer to Borroni and Zenga (2003).

The next section discusses the implementation of the tests based on the above considered indexes and the estimation of the corresponding power functions. Section 3 reports some results obtained by Monte Carlo simulations and shows how to choose the best testing procedure in different contexts. Section 4 summarizes and concludes.

2. Testing independence of two criteria

As above mentioned, the described rank correlation indexes can be used to test the independence of two criteria used to rank a sample of n units. To implement such tests, the values taken by each index need to be compared with the corresponding critical values, which can be determined by using the distribution of the indexes under the null hypothesis H_0 (that is independence of the two criteria). First of all, note that under H_0 every possible joint realization of the two sequences R_{11}, R_{21}, …, R_{n1} and R_{12}, R_{22},

…, R_{n2} have the same probability $1/(n!)^2$. The distribution of a test-statistic S can then be computed by enumerating all the $(n!)^2$ points of the sample space and by recording the frequency $f(s)$ of occurrence of each different value s taken by S; the probability of $S = s$ will then be given by the ratio $f(s)/(n!)^2$. Once the null distribution has been computed, the determination of the critical values of the corresponding test has to be faced. Indeed, H_0 can be tested against the lack of independence (two-sided alternative H_1), concordance (one-sided alternative H_1^+) or discordance (one-sided alternative H_1^-). If the alternative H_1 is considered, the null hypothesis has to be rejected whenever the generic rank correlation index is such that $S < s_1$ or $S > s_2$, where, under H_0, $\Pr\{S < s_1\} \le \alpha/2$ and $\Pr\{S > s_2\} \le \alpha/2$ (α being the significance level of the test). Note that, differently from ρ, τ and G, the null distribution of D is not symmetric and hence the critical values s_1 and s_2 need to be determined separately. In addition note that, as the distribution of the considered indexes is discrete, the nominal significance level α is rarely reached; hence a randomized test has often to be implemented. If one of the one-sided alternative is considered, the rejection rule consists of just one of the above inequalities; for instance, for H_1^+, one has to reject H_0 whenever the generic rank correlation index S is such that $S > s^+$ where, under H_0, $\Pr\{S > s^+\} \le \alpha$. A similar reasoning applies for H_1^-.

When the above considered rank correlation indexes are applied as test-statistics, the attention focuses on the comparison of the power functions of the corresponding test. As mentioned, the above described rank correlation indexes can be used to measure cograduation between two quantitative variables describing two populations. In this framework the use of nonparametric tests is justified if the usual assumption of normality on the two variables cannot be applied, as it often occurs in economic applications. In the following, we then aim to compare the performances of different tests based on rank correlation indexes, when the samples are drawn from known (possibly dependent) non-Normal populations. This comparison is accomplished by simulating the power functions of the tests by Monte Carlo methods. More specifically, after setting a known bivariate model for the two populations, a large number of bivariate samples are randomly drawn and the corresponding sequences of ranks are computed. Each test based on the above rank correlation indexes is then implemented and its power is estimated by computing the relative frequency of a correct rejection of the null hypothesis. By varying the extent of the dependence between the two populations, which should depend on a suitable parameter of the chosen model, and by iterating the above described procedure, the power function of each test is hence estimated.

The first problem to be faced is how to set a specific bivariate distribution model for the two populations. Of course, different choices can be made; however, note that a primary need of the study is to define a suitable parameter - accounting for the degree and the direction of dependence between the populations - which can be varied to simulate the power function of the tests. If two Normal populations were considered, a natural choice would be to set a bivariate Normal distribution with a varying coefficient of correlation. However, this choice does not reflect our need to consider departures from Normality. Hence a possible solution can be to choose distributions which can be considered as known "alterations" of the bivariate Normal model.

In this paper, the above conclusions are developed in two different ways. First, some suitable function of the bivariate Normal model, producing known bivariate distributions, are considered. Let (X, Y) be a bivariate standard Normal random variable, with correlation coefficient r; it is known that (e^X, e^Y) has a bivariate log-Normal distribution. Moreover, the transformations (X^2, Y^2) and ($|X|$, $|Y|$) generate a bivariate Chi-square and a bivariate half-Normal distribution respectively. The cited bivariate models are all characterized by non-Normal marginal distributions of the same kind, with a known extent of dependence (correlation) measurable by a suitable increasing function of $|r|$. Nevertheless, in real applications the two populations may show different characteristics. This fact can be considered in the simulated model by applying two different transformations to the components of (X, Y). However this may not produce an ordinary bivariate model, nor the two marginal distributions can be completely characterized.

For the above mentioned reasons, a second kind of "alteration" of a bivariate Normal distribution is applied in this paper. This method is proposed by Fleishman (1978) for the univariate case and by Vale and Maurelli (1983) for the multivariate case (see also Kotz et al. (2000)). Let (X, Y) be a standard Normal bivariate random variable with correlation coefficient r and let (X', Y') be a transformed random variable such that

$$X' = a_1 + b_1 X + c_1 X^2 + d_1 X^3 \qquad Y' = a_2 + b_2 Y + c_2 Y^2 + d_2 Y^3. \tag{6}$$

The resulting distribution of (X', Y') is not bivariate Normal. Moreover, the coefficients of the above transformations can be chosen so that the margins X' and Y' are characterized by a desired kind of departure from Normality. First of all, Vale and Maurelli (1983) propose to fix standardized margins. If, say, X' is considered, this choice obviously implies that $a_1 = -c_1$ and that b_1 can be determined in terms of c_1 and d_1; for instance, if c_1 and d_1 are set to 0, then $b_1 = \pm 1$. To determine the remaining values of the coefficients of the transformations, Vale and Maurelli (1983) propose to fix the level of skewness and kurtosis of the target margins, as measured by the indexes β_1 and β_2, which are known to be functions of the first four moments of a variable.[2] More specifically if, again, X' is chosen to have a standardized distribution with fixed values β_{11} and β_{21} for skewness and kurtosis, the coefficient of the transformation in (6) can be determined by setting $a_1 = -c_1$ and by solving the following system of equations for b_1, c_1 and d_1:

$$
\begin{aligned}
& b_1^2 + 6b_1d_1 + 2c_1^2 + 15d_1^2 - 1 = 0 \\
& 2c_1(b_1^2 + 24b_1d_1 + 105c_1^2 + 2) - \beta_{11} = 0 \\
& 24\left[b_1d_1 + c_1^2(1 + b_1^2 + 28b_1d_1) + d_1^2(12 + 48b_1d_1 + 141c_1^2 + 225d_1^2)\right] - \beta_{21} = 0
\end{aligned}
\qquad (7)
$$

(a similar reasoning applies for the coefficient of the transformation producing Y'). Furthermore, the correlation between X' and Y' can be easily shown to be an increasing function of $|r|$:

$$Corr(X', Y') = r(b_1b_2 + 3b_1d_2 + 3d_1b_2 + 9d_1d_2) + 2r^2c_1c_2 + 6r^3d_1d_2, \qquad (8)$$

so that the extent of dependence between X' and Y' can be varied according to the value of r. Note that, at the aim of this paper, the simulation of a large number of samples drawn from the populations described by (X', Y') is needed. The simplicity of the transformation generating this bivariate random variable, however, makes this task very easy.

The following section reports the results obtained by simulating the power functions of the considered tests in several different situations. We will be concerned just with the one-sided test of H_0 against the alternative of concordance H_1^+. The power function will be simulated by considering different "alterations" of the standard bivariate Normal model, producing ordinary bivariate distributions or other distributions with fixed levels of skewness and kurtosis, as described above.

3. Results of simulations

In this section some results obtained from several simulations, following the approach presented in the previous section, are reported. For each chosen distribution of the two populations, 100000 bivariate samples of a fixed size n were randomly drawn and the power of the one-sided tests (for H_0 against

[2] For a generic variable W with mean μ_W and variance σ^2_W, $\beta_1 = \sigma^3_W$ $E(W - \mu_W)^3$ and $\beta_2 = \sigma^4_W$ $E(W - \mu_W)^4$. Recall that, for a standardized Normal variable, $\beta_1 = 0$ and $\beta_2 = 3$. Note that sometimes the use of β_1 and β_2 as measures of skewness and kurtosis is considered as controversial (see Badaloni (1987) and Zenga (2005) for a discussion).

H_1^+) based on *D*, *τ*, *G* and *ρ* was estimated at different levels of dependence between the marginal distributions. Each test was implemented at a nominal significance level $\alpha = 0.05$ and randomized. The critical values of the considered tests were determined after computing the exact distribution of the corresponding test-statistics under H_0, as described in the previous section.

Table 1. *Bivariate Normal (left) and log-Normal (right) distribution (n = 7)*

D	*τ*	*G*	*ρ*	*r*	*D*	*τ*	*G*	*ρ*
0.0496	0.0496	0.0489	0.0494	0	0.0508	0.0505	0.0510	0.0504
0.0611	**0.0613**	0.0604	0.0608	0.05	*0.0612*	0.0608	0.0609	**0.0617**
0.0723	*0.0717*	0.0714	0.0723	0.10	**0.0713**	*0.0712*	0.0704	**0.0713**
0.0854	0.0845	0.0843	**0.0864**	0.15	**0.0871**	*0.0856*	0.0840	**0.0871**
0.1031	*0.1031*	0.1015	**0.1033**	0.20	*0.1023*	0.1022	0.1005	**0.1024**
0.1212	*0.1212*	0.1190	**0.1216**	0.25	*0.1204*	0.1199	0.1171	**0.1210**
0.1417	0.1415	0.1382	**0.1425**	0.30	*0.1431*	0.1426	0.1388	**0.1432**
0.1700	0.1679	0.1639	**0.1707**	0.35	*0.1699*	0.1686	0.1631	**0.1708**
0.1973	*0.1974*	0.1898	**0.1993**	0.40	*0.2003*	0.2000	0.1920	**0.2022**
0.2339	0.2331	0.2234	**0.2353**	0.45	*0.2329*	0.2306	0.2224	**0.2349**
0.2696	0.2689	0.2595	**0.2731**	0.50	*0.2718*	0.2704	0.2608	**0.2749**
0.3674	0.3645	0.3487	**0.3717**	0.60	*0.3659*	0.3656	0.3491	**0.3705**
0.4956	0.4917	0.4703	**0.5004**	0.70	*0.4921*	0.4912	0.4656	**0.4988**
0.6527	*0.6531*	0.6273	**0.6625**	0.80	*0.6526*	0.6520	0.6244	**0.6627**
0.8447	*0.8464*	0.8238	**0.8551**	0.90	0.8437	*0.8467*	0.8253	**0.8553**
1.0000	1.0000	1.0000	1.0000	1.00	1.0000	1.0000	1.0000	1.0000

At first let us analyze the results obtained for samples drawn from populations distributed according to ordinary bivariate models. In particular we focused on the following distributions: bivariate standard Normal (considered as a term of comparison), bivariate log-Normal, bivariate half-Normal and bivariate Chi-square. Different sample sizes, ranging from 5 to 10, were used. However, for the sake of simplicity, only the most significant results are presented here. Consider, for instance, Table 1, which regards the bivariate standard Normal distribution case with $n = 7$. Each row of Table 1 reports the estimated powers of the considered tests, corresponding to a correlation coefficient *r* ranging from 0 (first row) to +1 (last row). Notice that the first row of the table represents the actual significance level resulting from Monte Carlo simulations and that it is very close to the nominal value 0.05 for all the considered tests. Moreover, from Table 1 it seems clear that the test based on *ρ* behaves quite better than the other ones; this result is far from being unexpected since *ρ* can be seen as the nonparametric version of the correlation coefficient, which represents the parameter of the Normal bivariate model measuring dependence between the two marginal components. The test based on *G* is, at any level of the correlation between the margins, always the one with the lowest power. The test based on *D* can be seen as the “second best” after *ρ* and before *τ*, especially for low levels of correlation[3]; moreover, the same test is, in some cases, the direct competitor of the test based on ρ. It is worthwhile to note that it is desirable for a "good" test to show better performances for lower level of the dependence between the populations, that is for those situations which may be considered as more critical for they are characterized by a slight departure from the null hypothesis H_0. To appreciate such a characteristic, the estimated powers reported[4] in each row of Table 1 are obtained by increasing r of 0.05 when $r \leq 0.5$ and of 0.1 when $r > 0.5$.

Table 2. *Bivariate half-Normal distribution (left: n = 7, right: n = 9)*

D	τ	G	ρ	r	D	τ	G	ρ
0.0494	0.0499	0.0503	0.0498	0	0.0500	0.0497	0.0500	0.0505
0.0503	*0.0503*	**0.0504**	0.0498	0.05	*0.0513*	0.0505	0.0508	**0.0520**
0.0514	0.0513	**0.0519**	*0.0518*	0.10	**0.0521**	0.0515	0.0512	*0.0518*
0.0529	**0.0534**	*0.0529*	*0.0529*	0.15	*0.0541*	*0.0541*	*0.0541*	**0.0549**
0.0555	0.0553	0.0547	**0.0556**	0.20	**0.0568**	0.0564	0.0555	*0.0567*
0.0586	**0.0599**	0.0585	*0.0588*	0.25	**0.0611**	*0.0607*	0.0600	0.0606
0.0631	**0.0636**	0.0630	**0.0636**	0.30	0.0656	*0.0661*	0.0652	**0.0664**
0.0692	*0.0688*	0.0686	0.0685	0.35	0.0726	*0.0727*	0.0721	**0.0731**
0.0770	**0.0770**	0.0750	*0.0767*	0.40	*0.0832*	0.0826	0.0814	**0.0835**
0.0857	*0.0851*	0.0834	*0.0851*	0.45	0.0931	**0.0941**	0.0919	*0.0940*
0.0967	*0.0964*	0.0948	**0.0967**	0.50	0.1090	*0.1093*	0.1078	**0.1095**
0.1286	*0.1289*	0.1252	**0.1294**	0.60	*0.1551*	0.1546	0.1524	**0.1568**
0.1862	0.1844	0.1812	**0.1871**	0.70	*0.2317*	0.2313	0.2297	**0.2368**
0.2941	0.2932	0.2889	**0.2986**	0.80	0.3753	*0.3759*	0.3750	**0.3853**
0.5177	0.5154	0.5110	**0.5272**	0.90	0.6557	*0.6587*	0.6550	**0.6719**
1.0000	1.0000	1.0000	1.0000	1.00	1.0000	1.0000	1.0000	1.0000

Some further simulations, not reported here, show that the above conclusions hold similarly for different values of the sample size n. Table 1 (right) reports some results for the bivariate log-Normal case. Note that, to compare the performances of nonparametric tests of dependence, this model does not account for a real departure from Normality. Indeed in the log-Normal model, the variable X and Y producing (e^X, e^Y) are subjected to a monotone transformation under which the sequences of ranks are invariant. As a result, Table 1 (right) gives the same conclusions above reported for the bivariate Normal case. When the bivariate half-Normal model is considered, the test based on D often happens to be the best (especially for medium levels of correlation between the margins). However, the performance of the same test gets worse as the sample size increases; see Table 2.

[3] To give more readability, the best values of powers are reported in bold face, the second best in italic.

[4] This fact holds for all the following Tables. Notice that, for each considered model, the correlation between the margins turns out to be a different increasing function of r; in addition, the condition $r = 1$ does not necessarily imply a perfect correlation between the margins.

Table 3. *Bivariate Chi-square distribution (left: n = 7, right: n = 9)*

D	τ	G	ρ	r	D	τ	G	ρ
0.0492	0.0496	0.0497	0.0498	0	0.0499	0.0495	0.0494	0.0497
0.0509	0.0502	**0.0510**	**0.0510**	0.05	**0.0500**	0.0495	0.0497	*0.0499*
0.0505	**0.0510**	0.0504	**0.0510**	0.10	**0.0522**	*0.0517*	*0.0517*	**0.0522**
0.0539	0.0534	*0.0539*	**0.0544**	0.15	**0.0546**	*0.0543*	0.0542	**0.0546**
0.0558	0.0554	0.0547	*0.0555*	0.20	0.0558	*0.0562*	**0.0569**	0.0561
0.0579	*0.0577*	*0.0577*	**0.0579**	0.25	*0.0619*	0.0610	0.0605	**0.0620**
0.0639	**0.0642**	0.0626	**0.0642**	0.30	0.0664	*0.0666*	0.0660	**0.0668**
0.0691	0.0683	0.0673	*0.0688*	0.35	*0.0723*	0.0721	0.0716	**0.0725**
0.0747	**0.0751**	0.0727	*0.0748*	0.40	*0.0829*	0.0827	0.0816	**0.0834**
0.0839	0.0838	0.0824	**0.0843**	0.45	0.0948	**0.0954**	0.0930	*0.0952*
0.0974	*0.0975*	0.0940	**0.0979**	0.50	0.1084	*0.1090*	0.1056	**0.1092**
0.1298	*0.1298*	0.1274	**0.1302**	0.60	0.1531	*0.1536*	0.1519	**0.1556**
0.1881	0.1871	0.1838	**0.1895**	0.70	0.2321	*0.2329*	0.2299	**0.2374**
0.2905	0.2903	0.2865	**0.2947**	0.80	0.3803	*0.3804*	0.3772	**0.3894**
0.5161	0.5145	0.5069	**0.5245**	0.90	0.6591	*0.6624*	0.6575	**0.6740**
1.0000	1.0000	1.0000	1.0000	1.00	1.0000	1.0000	1.0000	1.0000

Concerning the bivariate Chi-square case (recall the above definition given in Section 2), reported in Table 3 for $n = 7$ and $n = 9$, it easily noted that the performance of the test based on D is worse than in the half-Normal case. Notice that, for both these bivariate models, the margins are characterized by strong asymmetry; however, when the marginals are Chi-square (with one degree of freedom), they shows heavier tails. It is fairly known that the distributions of the considered test-statistics do not depend on the marginal cdf of the populations. However, the reported results show that the presence of heavy tails in the observed sample can be considered as a guideline to reveal a bivariate model leading to a bad performance of D.

A second set of simulations was obtained by following the approach of Fleishman (1978) and Vale and Maurelli (1983), described in the previous section. Recall that this approach produces a wide range of bivariate distributions whose margins are of an unknown kind, with given levels of kurtosis and skewness; in addition, it allows to consider margins with the same or different distributions. Many situations were simulated; to organize the obtained results, consider the following numbered models, which will be used to identify the simulated bivariate distribution:

1. weak asymmetry and kurtosis index equals to 3;
2. strong asymmetry and kurtosis index equals to 3;
3. symmetry and kurtosis index lower than 3;
4. weak asymmetry and kurtosis index lower than 3;
5. symmetry and kurtosis index greater than 3;
6. weak asymmetry and kurtosis index greater than 3.

Table 4. *Combination Normal / model 1 and combination Normal / model 2 (n = 7)*

D	τ	G	ρ	r	D	τ	G	ρ
0.0497	0.0495	0.0490	0.0498	0	0.0495	0.0495	0.0499	0.0497
0.0612	**0.0612**	0.0602	*0.0607*	0.05	*0.0600*	0.0596	0.0596	**0.0602**
0.0720	**0.0722**	0.0715	**0.0722**	0.10	*0.0724*	**0.0725**	0.0718	**0.0725**
0.0851	0.0844	0.0844	**0.0860**	0.15	**0.0868**	0.0853	0.0851	*0.0867*
0.1038	0.1030	0.1012	*0.1036*	0.20	0.1016	**0.1025**	0.0997	*0.1020*
0.1214	0.1210	0.1187	**0.1215**	0.25	**0.1210**	*0.1204*	0.1176	**0.1210**
0.1419	0.1416	0.1377	**0.1426**	0.30	*0.1398*	0.1392	0.1355	**0.1406**
0.1698	0.1684	0.1641	**0.1705**	0.35	*0.1661*	0.1648	0.1606	**0.1668**
0.1976	0.1959	0.1912	**0.1989**	0.40	*0.1949*	0.1942	0.1886	**0.1972**
0.2335	*0.2338*	0.2227	**0.2347**	0.45	*0.2278*	0.2257	0.2179	**0.2292**
0.2698	0.2686	0.2591	**0.2730**	0.50	*0.2635*	0.2603	0.2525	**0.2660**
0.3679	0.3650	0.3483	**0.3713**	0.60	*0.3572*	0.3538	0.3409	**0.3618**
0.4952	0.4928	0.4702	**0.5013**	0.70	*0.4798*	0.4759	0.4577	**0.4852**
0.6536	0.6518	0.6266	**0.6622**	0.80	*0.6415*	0.6385	0.6151	**0.6521**
0.8450	*0.8459*	0.8240	**0.8550**	0.90	*0.8310*	0.8308	0.8105	**0.8419**
1.0000	1.0000	1.0000	1.0000	1.00	0.9964	**0.9980**	*0.9970*	0.9946

A first subset of simulations were obtained by combining one of the above listed model with the standard Normal distribution (recall that the standard Normal model is characterized by perfect symmetry and by a kurtosis index equal to 3). In these situations, the test based on D performs quite well, being often a good competitor of the test based on ρ and especially of the one based on τ. Conversely, the test based on G shows the worst performance. As an example, we report in Table 4 (left) the results obtained for the combination Normal / model 1, when $n = 7$. The good performance of the test based on D is emphasized when a standard normal margin is paired with a strong asymmetric model (see Table 4 (right) which shows the combination Normal / model 2, when $n = 7$).
When the assumption of Normality of one of the two margins is discarded, the performance of the test based on D is still good, at least if the same model is chosen for the two components. As an example, consider Table 5 (combination model 2 / model 2 and combination model 5 / model 5, both with $n = 7$). However, for bivariate distributions with different margins, the performances of the four tests do not show any particular tendency. Quite often the test based on ρ is the most powerful, but sometimes this fact happens to be true for D or τ. Concerning the comparison of D and τ as a second best test, sometimes D performs better than τ (see Table 6 (left) reporting combination model 3 / model 4, with $n = 7$); in other cases, the opposite situation is observed (Table 6 (right), combination model 5 / model 6, $n = 7$). Finally, Table 7 reports further comparison of models (combination model 4 / model 6 and combination model 3 / model 6, with $n = 7$).

Table 5. *Combination model 2 / model 2 and combination model 5 / model 5 (n = 7)*

D	*τ*	*G*	*ρ*	*r*	*D*	*τ*	*G*	*ρ*
0.0502	0.0494	0.0493	0.0499	0	0.0498	0.0498	0.0505	0.0505
0.0614	0.0607	0.0602	*0.0611*	0.05	**0.0618**	0.0610	0.0605	*0.0614*
0.0716	0.0712	0.0710	*0.0714*	0.10	**0.0731**	0.0722	0.0717	*0.0727*
0.0843	0.0834	0.0827	**0.0849**	0.15	**0.0871**	*0.0870*	0.0850	0.0867
0.1006	0.1003	0.0994	**0.1010**	0.20	*0.1023*	0.1013	0.0997	**0.1029**
0.1182	0.1178	0.1152	*0.1179*	0.25	*0.1220*	0.1203	0.1191	**0.1223**
0.1370	0.1360	0.1352	**0.1382**	0.30	*0.1450*	0.1440	0.1408	**0.1455**
0.1629	0.1600	0.1589	**0.1636**	0.35	*0.1690*	0.1673	0.1635	**0.1704**
0.1895	0.1883	0.1846	**0.1908**	0.40	*0.2000*	0.1979	0.1908	**0.2011**
0.2244	0.2227	0.2144	**0.2251**	0.45	*0.2313*	0.2308	0.2210	**0.2341**
0.2581	0.2544	0.2496	**0.2611**	0.50	*0.2706*	0.2686	0.2576	**0.2726**
0.3501	0.3458	0.3346	**0.3533**	0.60	*0.3660*	0.3646	0.3491	**0.3708**
0.4744	0.4708	0.4533	**0.4810**	0.70	*0.4956*	0.4946	0.4740	**0.5034**
0.6337	0.6318	0.6106	**0.6446**	0.80	*0.6543*	0.6541	0.6284	**0.6634**
0.8338	*0.8339*	0.8135	**0.8447**	0.90	0.8467	*0.8482*	0.8251	**0.8556**
1.0000	1.0000	1.0000	1.0000	1.00	1.0000	1.0000	1.0000	1.0000

Table 6. *Combination model 3 / model 4 and combination model 5 / model 6 (n = 7).*

D	*τ*	*G*	*ρ*	*r*	*D*	*τ*	*G*	*ρ*
0.0499	0.0498	0.0494	0.0496	0	0.0495	0.0497	0.0497	0.0497
0.0588	**0.0594**	0.0584	*0.0588*	0.05	0.0597	0.0589	*0.0598*	**0.0600**
0.0708	*0.0713*	0.0707	**0.0718**	0.10	**0.0733**	*0.0732*	0.0722	0.0730
0.0844	*0.0845*	0.0832	**0.0846**	0.15	*0.0862*	0.0856	0.0838	**0.0863**
0.1025	0.1019	0.1001	*0.1024*	0.20	*0.1014*	0.1012	0.0990	**0.1019**
0.1220	0.1211	0.1183	*0.1218*	0.25	0.1216	*0.1219*	0.1186	**0.1230**
0.1460	0.1441	0.1409	**0.1465**	0.30	*0.1448*	0.1437	0.1412	**0.1459**
0.1688	0.1675	0.1622	**0.1690**	0.35	*0.1707*	0.1697	0.1651	**0.1727**
0.1984	0.1969	0.1910	**0.2000**	0.40	*0.1988*	0.1985	0.1918	**0.2006**
0.2293	0.2272	0.2189	**0.2306**	0.45	*0.2326*	0.2312	0.2244	**0.2341**
0.2675	0.2652	0.2569	**0.2703**	0.50	*0.2710*	0.2687	0.2579	**0.2745**
0.3627	0.3595	0.3456	**0.3662**	0.60	*0.3645*	0.3638	0.3462	**0.3690**
0.4849	0.4821	0.4607	**0.4917**	0.70	*0.4932*	0.4904	0.4694	**0.4993**
0.6476	0.6456	0.6204	**0.6579**	0.80	*0.6521*	0.6517	0.6260	**0.6623**
0.8395	*0.8409*	0.8199	**0.8506**	0.90	0.8448	*0.8468*	0.8230	**0.8547**
0.9988	**0.9994**	*0.9989*	0.9982	1.00	1.0000	1.0000	1.0000	1.0000

Table 7. *Combination model 4 / model 6 and combination model 3 / model 6 (n = 7)*

D	*τ*	*G*	*ρ*	*r*	*D*	*τ*	*G*	*ρ*
0.0498	0.0501	0.0495	0.0497	0	0.0507	0.0506	0.0506	0.0511
0.0599	**0.0602**	0.0592	*0.0599*	0.05	0.0589	**0.0595**	*0.0593*	*0.0593*
0.0716	**0.0718**	0.0707	0.0715	0.10	**0.0729**	*0.0723*	0.0713	*0.0723*
0.0873	*0.0869*	0.0853	0.0867	0.15	**0.0871**	0.0856	0.0853	*0.0865*
0.1028	**0.1028**	0.1005	*0.1027*	0.20	0.1033	*0.1034*	0.1004	**0.1035**
0.1203	*0.1206*	0.1174	**0.1213**	0.25	**0.1225**	0.1216	0.1194	*0.1224*
0.1436	0.1427	0.1384	**0.1444**	0.30	*0.1436*	0.1434	0.1389	**0.1447**
0.1687	0.1678	0.1633	**0.1704**	0.35	*0.1680*	0.1671	0.1615	**0.1686**
0.1977	0.1959	0.1901	**0.1990**	0.40	*0.1995*	0.1974	0.1921	**0.2010**
0.2299	0.2284	0.2202	**0.2312**	0.45	*0.2317*	0.2308	0.2226	**0.2337**
0.2690	0.2664	0.2563	**0.2716**	0.50	*0.2704*	0.2678	0.2576	**0.2722**
0.3642	0.3624	0.3465	**0.3681**	0.60	*0.3683*	0.3658	0.3496	**0.3715**
0.4867	*0.4872*	0.4640	**0.4955**	0.70	*0.4929*	0.4913	0.4688	**0.5004**
0.6485	0.6484	0.6209	**0.6577**	0.80	0.6533	*0.6543*	0.6266	**0.6627**
0.8423	*0.8426*	0.8217	**0.8518**	0.90	0.8443	*0.8465*	0.8237	**0.8552**
0.9988	**0.9993**	0.9987	0.9981	1.00	1.0000	1.0000	1.0000	1.0000

4. Summary and conclusions

In this paper the performance of a new nonparametric test for the independence of two criteria is discussed. This test is based on Gini's mean difference computed on the total ranks assigned to each sampled unit according to the chosen criteria of sorting. The performance of the test is measured by simulating its power function via Monte Carlo methods when it is applied as a one-sided test of independence against concordance.

More specifically, it is first assumed that the two criteria of ranking are based on the values taken by two quantitative variables on each sample unit; a specific bivariate model, which can guarantee the existence of a parameter measuring dependence between the two components, is then set. The choice of the bivariate model reflects the common situation of sampling from non-Normal populations, usually faced in economic applications. A large number of bivariate samples are then randomly drawn from the chosen model. The power function of the test is hence estimated by computing the relative frequency of a correct rejection of the null hypothesis and by varying the value of the parameter measuring dependence. The considered test is then compared with other common nonparametric tests based on Spearman's rho, Kendall's tau and Gini's cograduation index.

The Monte Carlo simulations reported in this paper show that, among those which are examined, none of the chosen tests can be considered uniformly as the best. However, the test based on ρ very often performs better than the other tests. This fact can be considered fairly expected when the bivariate Normal model is chosen for the two populations, but it is not trivial when the populations are far from Normality. However, as the departure from Normality becomes stronger, the tests based on D and on τ can be seen as more direct competitors of the one based on ρ.

To get into details, the test based on D seems to give the best results, being often more powerful than the other considered tests, in such situations which can be identified as sampling from asymmetric but not heavy-tailed populations. In addition, the same test has a good performance when the level of dependence of the two populations is low; this can be considered as a good property of a test, as situations of a low dependence are usually hard to be detected. Conversely, the performance of the test based on D seems to get worse as the sample size increases and when the two marginal populations are not equally distributed. As a final remark, it is important to stress that in all the analyzed cases the test based on D results to be more powerful than the one based on G, which is always the worst test.

Of course the reported set of simulations is far from being complete. First of all, recall that only the one-sided version of the considered tests of independence against concordance has been simulated. Furthermore, no work has been done to compare the performance of the tests when the criteria of sorting are not based on the values taken by quantitative variables.

It has to be pointed out, indeed, that this work essentially deals with a structure of dependence based on correlation: the varying parameter in the simulations of powers is always an increasing function of the coefficient of correlations between the marginal variables. In addition, even in this framework, different parametric models could be selected for the two variables used to sort the sampled units. It is believed, however, that the reported results can be considered as important guidelines for an accurate use of the test based on *D*, especially in economic applications where Gini's mean difference has been widely appreciated as a measure of concentration and dispersion.

References

Badaloni, M. (1987) "Skewness and Abnormalities of Statistical Distributions", in: Naddeo, A. (Ed.), *Italian Contributions to the Methodology of Statistics*, Cleup, Padova, 603-612.

Borroni, C. G., Zenga, M. (2003) "A test of Concordance Based on the Distributive Compensation Ratio", *Rapporti di Ricerca del Dipartimento di Metodi Quantitativi per l'Economia - Università degli Studi di Milano Bicocca*, 51.

Cifarelli, D. M., Conti, P. L., Regazzini, E. (1996) "On the Asymptotic Distribution of a General Measure of Monotone Dependence", *Annals of Statistics*, **24**, 1386-1399.

David, H. A. (1968) "Gini's Mean Difference Rediscovered", *Biometrika*, **55**, 573-575.

Fleishman, A. I. (1978) "A Method for Simulating Non-Normal Distributions", *Psychometrika*, **43**, 521-532.

Gini, C. (1912) "Variabilità e Mutabilità; Contributo allo Studio delle Distribuzioni e Relazioni Statistiche", *Studi Economico - Giuridici, R. Università di Cagliari.*

Gini, C. (1954) *Corso di Statistica*, Veschi, Rome.

Kendall, M., Gibbons, J. D. (1990) *Rank Correlation Methods*, Oxford University Press, New York.

Kotz, S., Johnson, N. L. (1982) *Encyclopedia of Statistical Sciences*, Wiley, New York.

Kotz, S., Balakrishnan, N., Johnson, N. (2000) *Continuous Multivariate Distributions. Volume 1: Models and Applications*, Wiley, New York.

Vale, C. D., Maurelli, V. A. (1983) "Simulating Multivariate Non-Normal Distributions", *Psychometrika*, **48**, 465-471.

Zenga, M. (2003) "Distributive Compensation Ratio Derived from the Decomposition of the Mean Difference of a Sum", *Statistica & Applicazioni*, **1**, 19-27.

Zenga, M. (2005) "Kurtosis", in: Kotz, S., Balakrishnan, N., Read, C. B, Vidakovic, B., Johnson, N. L., *Encyclopedia of Statistical Sciences - Second Edition*, Wiley, New York.