Statistica Sinica **10**(2000), 577-593

RANK TESTS FOR INDEPENDENCE — WITH A WEIGHTED CONTAMINATION ALTERNATIVE

Grace S. Shieh, Zhidong Bai and Wei-Yann Tsai

Academia Sinica, National Sun-Yat-Sen University and Columbia University

Abstract: Two rank tests for independence of bivariate random variables against an alternative model with weighted contamination are proposed. The model may emphasize the association of X and Y on items with high ranks in one variable (say X) and generalizes an alternative in Hájek and Šidák (1967). The model may be applied to both complete paired data and paired data which is *truncated* in one variable. We derive the locally most powerful rank (LMPR) test under the alternative setting. The proposed tests turn out to be asymptotic LMPR tests under Logistic and Extreme Value families. Under the null hypothesis of independence, both rank statistics have limiting normal distributions. An application to a data set from a special education program in Taiwan and a simulation study are presented. We also apply the Shapiro-Francia test to find the minimum sample sizes for approximate normality of exact distributions of the proposed test statistics.

Key words and phrases: Association, independence test, Kendall's Tau, rank, Spearman's Rho.

1. Introduction

Let $\{(X_i, Y_i),\ 1 \le i \le n\}$ be an independently and identically distributed (i.i.d.) sample from a bivariate population (X, Y). We introduce two rank statistics for testing independence of X and Y against an alternative with weighted contamination as follows:

$$X_i = X_i^* + w(X_i^*)\Delta Z_i \text{ and } Y_i = Y_i^* + \Delta Z_i, \quad 1 \le i \le n, \tag{1}$$

where X_i^*, Y_i^* and Z_i are mutually independent random variables (r.v.'s); Δ a constant and $w(x)$ monotone in x. Under (1) it is clear that if $\Delta = 0$, X and Y are independent, and the larger Δ is, the more dependent are X and Y. Thus the constant Δ may be regarded as a dependence or mixing coefficient. For more details, see Section 2.1.

Measures of linear association were discussed and studied around 1900 (Galton (1877), Pearson (1896), Spearman (1904) and Kendall (1938)). In the last two decades, measures of weighted correlation have been extensively discussed in Salama and Quade (1982), Iman and Conover (1987), and others. For details, see the thorough review paper by Quade and Salama (1992). Motivated by applications in sensitivity analysis, Iman and Conover (1987) proposed a measure, the Pearson correlation coefficient computed on Savage scores (Savage (1956)). This measure reflects well the importance of agreement on the top ranks.

The alternative in (1) may stress the association of X and Y on items with top ranks in one variable (say X) and generalizes the alternative to independence in Hájek and Šidák (1967) by introducing a weight function $w(\cdot)$. The proposed alternative model allows our rank tests to be applied to both complete and *truncated* data.

Examples of the proposed model in (1) are frequently encountered in real life. Besides complete data sets, the new model may be adopted in the following situations. For saving costs and/or time, one may exclude many subjects or items with bottom ranks on one feature in a screening procedure, and then focus on examining those which passed the screening. For example, a recruiting committee may screen the applicants by their resumes first and interview only a few candidates. In this case, one r.v. involved (say X) could be the applicants' qualifications shown in the resume and the other r.v. (say Y) is the applicants' qualifications evaluated from an interview. Another example occurs in education of the gifted. In an identification procedure of the gifted in natural sciences, suppose that students will take both Mathematics and Physics aptitude tests. A common belief is that the test scores are positively correlated. Therefore, to save costs, one may test the students on one subject, say Mathematics, first and then further test the top-ranked (say the top 10%) in Physics. Here X_i is the ith student's Mathematics aptitude score and Y_i the Physics aptitude score. Note that all values of X_i's are observed but we observe only the values of Y whose corresponding values of X are top-ranked.

In the above cases, the procedures are fair provided the random variables are positively correlated or dependent. Thus testing the independence of X and Y against the alternative in (1) is an important issue. The rank tests are good tools for this testing problem and this is confirmed by our results and a simulation study.

In Section 2, we first derive a general form of the locally most powerful rank (LMPR) test under (1). We then show that the two proposed tests are asymptotic LMPR tests under Logistic and Extreme Value families. Further, their limiting distributions under the null hypothesis are derived. In Section 3, an application to a data set from a special education program in Taiwan is presented. Section 4 includes a power comparison of the new tests to those studied in Iman and Conover (1987) and Shieh (1998). Powers of the tests are summarized in Tables 1-4. The minimum sample sizes for approximate normality of exact distributions of both tests are also studied. We conclude with some remarks in Section 5.

Some critical values of the tests can be obtained from the corresponding author upon request.

2. New Rank Tests

Recall that (X_i, Y_i), $1 \le i \le n$, are i.i.d. bivariate r.v.'s. Let $(i, R_i), i = 1, \ldots, n$, be paired rankings, where R_i is the rank of Y whose corresponding X has rank i among $\{X_j\}$. We assume that there are no ties among the variables being ranked. Iman and Conover (1987) propose the top-down correlation coefficient

$$r_T = (\sum_{i=1}^{n} S_i S_{R_i} - n)/(n - S_1), \tag{2}$$

where the S_i are Savage scores defined as

$$S_i = \sum_{j=i}^{n} 1/j. \tag{3}$$

The correlation coefficient r_T reflects the association of top ranks well, and is the LMPR test statistic under the alternative to independence on page 75 of Hájek and Šidák (1967) when both X and Y have extreme value distributions. However, r_T is well defined only when the samples are fully observed. As mentioned earlier, in many cases the data involved are truncated. In addition, a weighting structure is needed when one wants to emphasize certain part(s) of the ranks, say those in the middle. To include various types of weights and to address the weighting structure issue in the alternative, we propose the following rank statistics: weighted Spearman's Rho (T_s) and weighted top-down statistic (T_t).

With objectively or subjectively chosen weights w_i that depend solely on i, a weighted Spearman's Rho is defined by

$$T_s = \sum_{i=1}^{n} w_i(i - (n+1)/2)(R_i - (n+1)/2) \tag{4}$$

and a weighted top-down statistic is

$$T_t = \sum_{i=1}^{n} w_i(S_{n-i+1} - 1)(S_{n-R_i+1} - 1).$$

For instance, we can take $w_i = I_{[i \le m]}$, where $m = [(n+1)p]$ and $0 < p \le 1$ is roughly the percentage of the observed items (subjects). In general, we choose p to have small loss in significance level (P-value) and to save computation. This is illustrated further in the gifted students example of Section 4.

With equal weights $w_i \equiv 1/(n - S_1)$, T_t reduces to r_T. Note that r_T emphasizes agreement of the top ranks by substituting Savage scores for ranks into the

Pearson correlation coefficient, while T_t puts additional weights on top-ranked ones.

Remark 1. With equal weights $12/[n(n^2-1)]$, T_s reduces to the Spearman's Rho. Thus for convenience, one should not assume that the weights sum up to one. Instead of (4), we can write T_s as $\sum_i \{w_i[i-(n+1)/2]-c_n\}[R_i-(n+1)/2]$, where $c_n = \sum_i w_i[i-(n+1)/2]/(\sum_i w_i)$, which is a centered version. Likewise, a centered version of T_t can be defined.

Remark 2. The dependence of w_i on i, and hence on $\{X_i\}$, implies that the truncation depends on the ranks of $\{X_i\}$. Thus in practice we let X be the variable that can be, or is, easily truncated.

We note that the weights in T_s and T_t can be adjusted easily to test for both top-down and bottom-up correlation alternatives. Recall that the top-down correlation emphasizes the agreement in the top ranks, whereas the bottom-up correlation stresses the agreement in the bottom ranks.

2.1. LMPR tests

In the following, we first derive a general form of the LMPR test under the weighted contamination alternative in (1). Further, we show that T_s and T_t are the asymptotic LMPR tests with respect to Logistic and Extreme Value families, respectively. Recall (1). Usually, the weight function $w(x)$ is increasing in x, and in many cases it is also differentiable. However these are not essential in our limit theorems. The alternative hypothesis of a weighted contamination can be detected by either T_s or T_t. The weighted rank tests are especially useful when the marginal distributions of the variables being ranked are skewed to the right.

Let X^* and Y^* have densities $f(x)$ and $g(y)$, respectively, while the distribution of Z_i is arbitrary. For ease of statement, in the sequel we assume that $w(x)$ is increasing and differentiable. Thus for given x and Δz, the equation $x = x^* + w(x^*)\Delta z$ has a unique solution for x^*, denoted $x^* = s(x, \Delta z)$. Then the i.i.d. sample (X_i, Y_i), $i = 1, \ldots, n$, has a density given by $q_\Delta = \prod_{i=1}^n h_\Delta(x_i, y_i), -\infty < \Delta < \infty$, where

$$h_\Delta(x,y) = \int_{-\infty}^{\infty} \frac{f(x^*)g(y-\Delta z)}{1+w'(s(x,\Delta z))\Delta z} dM(z),$$

and $M(z)$ is a distribution of Z with mean μ_z and finite variance σ_z^2.

Let $X_{(i)}$ and $Y_{(i)}$ be the ith order statistics of $\{X_i\}$ and $\{Y_i\}$, respectively. Further, let $a_n(i, w, f) = E\{-[(wf)'/f](X_{(i)})\}$ and $b_n(i, g) = E\{-[g'/g](Y_{(i)})\}$ denote the score functions corresponding to the density f and weight function w, and to density g, respectively. The following theorem states the general form of the LMPR test.

Theorem 1. *Assume that* $\int_{-\infty}^{\infty} |(wf)'(x)|dx < \infty$, $\int_{-\infty}^{\infty} |g'(x)|dx < \infty$, *and that* $(wf)'(x)$ *and* $g'(x)$ *are continuous almost everywhere. Then the test with critical region* $\sum_{i=1}^{n} a_n(i,w,f)b_n(R_i,g) \geq k$ *is the LMPR test for* $H_0 : \Delta = 0$ *against* $H_1 : \Delta > 0$.

The proof is given in Appendix 1.

Remark 3. When $w(x)$ is not continuously differentiable (or even continuous) but $w(x)f(x)$ is of bounded variation, if we define the score function $a_n(i,w,f)$ as

$$a_n(i,w,f) = \frac{n!}{(i-1)!(n-i)!} \int F^{i-1}(x)(1-F(x))^{n-i} d(w(x)f(x)), \qquad (5)$$

then Theorem 1 remains valid.

For $w(x) = I_{[0,p]}(F(x)) = I_{(-\infty,\xi_p]}(x)$, where $\xi_p = F^{-1}(p)$, we have from (5) that

$$\begin{aligned} &a_n(i,w,f) \\ &= \frac{n!}{(i-1)!(n-i)!}\left[\int_{-\infty}^{\xi_p} F^{i-1}(x)(1-F(x))^{n-i} f'(x)dx - p^{i-1}(1-p)^{n-i} f(\xi_p)\right] \\ &= E[I[U_{(i)} \leq p] \cdot \varphi(U_{(i)},f)] - \frac{n!}{(i-1)!(n-i)!} p^{i-1}(1-p)^{n-i} f(\xi_p), \end{aligned}$$

where $U_{(i)}$ is the ith ordered sample from $U[0,1]$ and $\varphi(u,f) = (f'/f)(F^{-1}(u))$. Then in the LMPR test statistic the factor $a_n(i,w,f)$ can be approximated by

$$a_n(i,w,f) \approx I_{[i/(n+1) \leq p]} \varphi(i/(n+1),f).$$

The reason is the following: for $|np-i| \geq c\sqrt{n}$ with a large constant c, $E[I_{[U_{(i)} \leq p]}\varphi(U_{(i)},f)] \approx I_{[i/(n+1)\leq p]}\varphi(i/(n+1),f)$, and there are only $[2c\sqrt{n}]$ negligible terms satisfying $|np-i| < c\sqrt{n}$.

Corollary 1. *If* F *and* G *are from the Logistic family, then the test* T_s *with* $w_i = I_{[i\leq m]}$, $m = [(n+1)p]$ *and critical region* $T_s \geq k$, *where* k *is a constant, is the asymptotic LMPR test for* $H_0 : \Delta = 0$ *versus* $H_1 : \Delta > 0$ *at* (1).

Proof. By (13) in page 67 of Hájek and Šidák (1967), $b_n(i,f) \cong \varphi(i/(n+1),f)$ is the approximate scores corresponding to f, and when f is logistic, $\varphi(i,f) = 2i-1$. Thus $\varphi(i/(n+1),f) \cong 2i/(n+1) - 1 \propto i - (n+1)/2$ and by (5), $a_n(i,w,f) \cong I_{[i/(n+1)\leq p]}\varphi(i/(n+1),f)$. It follows that the asymptotic LMPR test statistic is proportional to $\sum_{i=1}^{n} w_i[i-(n+1)/2][R_i-(n+1)/2]$.

Similarly, for the Extreme Value family with p.d.f. $f(x) = exp\{x - e^{-x}\}$, we have $\varphi(i,f) = -ln(1-i)-1$. Taking $\varphi(i/n,f)$ as approximate scores corresponding to f and by $ln(i/n) \cong -S_{i+1}$, we obtain $\varphi(i/n,f) \cong -ln((n-i)/n) - 1 \cong$

$S_{n-i+1} - 1$. Again by (5), we have $a_n(i, w, f) \cong I_{[1/(n+1) \leq p]}(S_{n-i+1} - 1)$. This and the fact that $b_n(R_i, g) \cong S_{n-R_i+1} - 1$ imply Corollary 2.

Corollary 2. *If $F(x)$ and $G(x)$ are from the Extreme Value family, then the test T_t with $w_i = I_{[i \leq m]}$, $m = [(n+1)p]$ and critical region $T_t \geq k$, where k is a constant, is the asymptotic LMPR test for $H_0 : \Delta = 0$ versus $H_1 : \ \Delta > 0$ at* (1).

2.2. Null limit distributions

In this section, the asymptotic distributions of T_s and T_t are derived under H_0, the hypothesis of independence. Let $I(f)$ denote Fisher information, $I(f) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 f(x)\, dx$.

Theorem 2. *Assume that H_0 holds, $I(f) < \infty$ and $I(g) < \infty$. Then $\sqrt{T}_s / \{n(n^2 - 1)p[3(1-p)^3 + p^2]/12\}^{1/2} \rightarrow_D N(0, 1)$.*

The proof is given in Appendix 2. When $p = 1$, T_s reduces to the usual Spearman's Rho r_s and Var$(\sqrt{T}_s)$ equals $[n(n^2 - 1)]/12$.

Theorem 3. *Assume H_0 holds, $I(f) < \infty$ and $I(g) < \infty$. Then $T_w/(nc_p)^{1/2} \rightarrow_D N(0, 1)$, where $c_p = p - (1-p)(2-p)ln^2(1-p)$.*

The proof is given in Appendix 3. When $p = 1$, after some normalization T_t reduces to r_T as in Iman and Conover (1987), and $c_p = 1$ implies that Var$(T_t) = n(1 + o(1))$ which agrees with the fact that Var$(r_T) = n - S_1$.

3. Simulation Results

In this section, the results of a power comparison and the minimum n required for approximate normality of exact distributions of the new tests are presented.

3.1. Power comparison

We first compare the powers of the new tests with those of the top-down statistic r_T in (2) (Iman and Conover (1987)) and the weighted Kendall's Tau in Shieh (1998) for finite sample sizes. A weighted Kendall's Tau is defined as

$$\tau_w = \frac{1}{\sum_{i \neq j} w_i w_j} \sum_{i \neq j} w_i w_j sgn(i - j) sgn(R_i - R_j),$$

where $w_i = I_{[i \leq m]}$ and $m = [(n+1)p]$.

The alternative used is at (1). Four choices of p were studied, namely $p = 0.1$, 0.2, 0.3 and 0.5. Let ρ denote the correlation between X_i and Y_i. The relationship of ρ versus Δ is: $\rho = p\Delta^2\sigma_z^2/[(\sigma_x^2 + p\Delta^2\sigma_z^2)(\sigma_y^2 + \Delta^2\sigma_z^2)]^{1/2}$. The choices of ρ include $\rho = 0.0$ (which gives the null hypothesis), $\rho = 0.1(0.1 \text{ or } 0.2)\rho_{max}$, where ρ_{max} is $\sqrt{p}$ rounded to the first decimal place.

Table 1. Powers of weighted Spearman's Rho, Kendall's Tau and top-down statistic with $p = 0.1$.

			N(0,1)				Logistic(0,−1)				Extreme value dist.[a]			
$n+1$	m	ρ	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T
10	1	.1	.055	—	.063	.057	.061	—	.069	.065	.052	—	.056	.063
		.2	.085	—	.110	.075	.082	—	.106	.076	.079	—	.100	.079
		.3	.100	—	.143	.083	.103	—	.145	.079	.100	—	.146	.081
20	2	.1	.070	.079	.112	.104	.080	.076	.128	.097	.053	.073	.066	.092
		.2	.147	.089	.304	.108	.136	.086	.286	.115	.144	.085	.277	.117
		.3	.260	.098	.492	.131	.219	.096	.489	.121	.264	.095	.501	.128
30	3	.1	.087	.144	.123	.109	.084	.142	.135	.104	.072	.133	.076	.099
		.2	.220	.205	.401	.154	.236	.204	.384	.144	.209	.193	.344	.142
		.3	.435	.271	.653	.158	.454	.269	.662	.170	.449	.264	.650	.177
50	5	.1	.135	.378	.167	.142	.123	.373	.202	.134	.080	.329	.080	.126
		.2	.426	.699	.595	.221	.387	.676	.569	.199	.333	.618	.493	.198
		.3	.745	.955	.867	.256	.724	.946	.849	.254	.759	.933	.845	.266
100	10	.1	.233	.857	.284	.227	.227	.754	.323	.203	.117	.670	.093	.182
		.2	.756	.989	.869	.340	.675	.980	.836	.346	.587	.961	.750	.338
		.3	.982	1.000	.993	.424	.980	1.000	.991	.430	.989	.999	.986	.454
200	20	.1	.422	.987	.466	.383	.349	.966	.508	.345	.183	.931	.137	.295
		.2	.969	1.000	.990	.583	.936	1.000	.983	.543	.872	.999	.947	.537
		.3	1.000	1.000	1.000	.704	1.000	1.000	1.000	.702	1.000	1.000	1.000	.730
400	40	.1	.664	1.000	.731	.574	.592	1.000	.744	.525	.253	.999	.194	.452
		.2	1.000	1.000	1.000	.860	.999	1.000	1.000	.825	.989	1.000	.997	.813
		.3	1.000	1.000	1.000	.954	1.000	1.000	1.000	.943	1.000	1.000	1.000	.956

[a] p.d.f. $f(x) \sim exp\{x - exp\{x\}\}$.

The powers of T_s, τ_w and T_t under the alternative, with X_i, Y_i and $Z_i \sim N(0,1)$, X_i, Y_i and $Z_i \sim Logistic(0,-1)$, and X_i, Y_i and $Z_i \sim$ the Extreme Value distribution with p.d.f. $f(x) \sim exp\{x - e^x\}$, are summarized in Tables 1-4, respectively. In each simulation, the number of replications used was 5,000 which yields a standard error of about .0071. The sample sizes studied are 9, 19, 29, 49, 99, 199 and 399. The 95% confidence interval for a power is $(\text{power} - 1.96 \times 0.0071, \text{power} + 1.96 \times 0.0071)$. We note that for fixed n and ρ, as p increases the power of r_T increases, although r_T does not depend on p. (Since a larger value of p leads to a larger "top-down" correlation of X and Y, it leads to a significantly larger r_T value.) For instance, for $n+1 = 10$, $\rho = 0.3$, the power of r_T in Table 2 ($p = 0.2$) is significantly larger than that in Table 1 ($p = 0.1$). The test statistics are discrete and we have randomized so that empirical powers of the tests with $\rho = 0$ under H_1 are equal to 0.05. For instance, for tests with

$n+1=10$ and $p=0.1$, the largest and second largest values of T_t equal to 16 and 12 with empirical probabilities (say) p_1 and p_2 respectively, and we added the empirical power of $T_t=12$ with weight $(0.05-p_1)/p_2$ to that of $T_t=16$ to obtain 0.05.

Table 2. Powers of weighted Spearman's Rho, Kendall's Tau and top-down statistic with $p=0.2$.

			N(0,1)				Logistic(0,−1)				Extreme value dist.			
$n+1$	m	ρ	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T
10	2	.1	.060	.067	.064	.079	.064	.069	.074	.080	.050	.067	.055	.072
		.2	.112	.077	.164	.099	.134	.079	.183	.113	.085	.077	.129	.107
		.3	.174	.086	.292	.130	.173	.086	.286	.123	.159	.084	.275	.122
		.4	.265	.092	.445	.162	.258	.094	.419	.150	.263	.094	.444	.155
20	4	.1	.102	.169	.084	.109	.095	.157	.097	.103	.074	.136	.055	.094
		.2	.218	.291	.252	.166	.212	.269	.266	.161	.150	.243	.165	.153
		.3	.411	.441	.499	.228	.387	.423	.486	.218	.344	.375	.437	.216
		.4	.615	.678	.719	.258	.605	.653	.709	.245	.631	.613	.731	.255
30	6	.1	.124	.262	.093	.136	.120	.267	.112	.143	.090	.226	.067	.126
		.2	.295	.498	.347	.209	.281	.491	.351	.206	.209	.427	.218	.191
		.3	.564	.729	.653	.277	.522	.720	.621	.275	.484	.658	.569	.273
		.4	.836	.941	.895	.348	.813	.926	.874	.346	.831	.900	.879	.363
50	10	.1	.150	.472	.116	.180	.162	.431	.166	.169	.101	.367	.074	.151
		.2	.429	.824	.490	.312	.436	.827	.511	.303	.297	.763	.304	.279
		.3	.790	.971	.853	.429	.730	.950	.822	.389	.671	.913	.756	.380
		.4	.974	1.000	.988	.536	.959	.998	.982	.508	.973	.996	.980	.528
100	20	.1	.256	.764	.188	.300	.263	.717	.256	.274	.148	.625	.096	.228
		.2	.695	.989	.746	.517	.680	.979	.747	.476	.475	.948	.451	.439
		.3	.969	1.000	.985	.678	.960	.999	.983	.638	.929	.998	.956	.630
		.4	1.000	1.000	1.000	.822	.999	1.000	.999	.809	1.000	1.000	1.000	.828
200	40	.1	.383	.973	.287	.443	.420	.954	.393	.432	.224	.920	.126	.372
		.2	.942	.999	.955	.794	.926	1.000	.954	.754	.747	.999	.715	.689
		.3	1.000	1.000	1.000	.921	.999	1.000	1.000	.910	.997	1.000	.999	.901
		.4	1.000	1.000	1.000	.978	1.000	1.000	1.000	.978	1.000	1.000	1.000	.984
400	80	.1	.665	.999	.499	.710	.670	.999	.621	.658	.360	.994	.193	.564
		.2	.997	1.000	.999	.964	.997	1.000	.998	.946	.952	1.000	.933	.916
		.3	1.000	1.000	1.000	.997	1.000	1.000	1.000	.996	1.000	1.000	1.000	.995
		.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

From Tables 1-4, we find that for values of samples sizes and p studied, all three rank tests, T_s, τ_w and T_t, are much more powerful than r_T. When $m \le 4$, T_t has slightly more power than T_s and τ_w, whereas for $m \ge 5$, T_s and τ_w perform better than T_t in most cases. Although τ_w is not the LMPR test with respect to either the Logistic or the Extreme Value distribution, it does perform rather well. It would be interesting to investigate under what distribution τ_w or $\hat{\tau}_w$ (the projection of τ_w onto the linear rank statistics) is the LMPR test.

Table 3. Powers of weighted Spearman's Rho, Kendall's Tau and top-down statistic with $p = 0.3$

			N(0, 1)				Logistic(0, −1)				Extreme value dist.			
$n+1$	m	ρ	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T
10	3	.1	.087	.086	.074	.087	.088	.084	.083	.083	.072	.081	.064	.077
		.3	.271	.152	.282	.158	.264	.145	.290	.149	.233	.138	.246	.144
		.5	.545	.235	.591	.223	.522	.230	.601	.233	.548	.224	.617	.242
20	6	.1	.122	.180	.093	.123	.126	.193	.118	.114	.096	.173	.069	.102
		.3	.463	.548	.483	.270	.450	.514	.489	.251	.382	.457	.399	.238
		.5	.880	.939	.908	.423	.883	.918	.898	.418	.902	.894	.902	.438
30	9	.1	.146	.271	.105	.146	.169	.271	.137	.157	.116	.236	.076	.140
		.3	.613	.794	.639	.344	.594	.754	.627	.348	.497	.682	.508	.331
		.5	.978	.997	.986	.562	.964	.994	.972	.555	.976	.988	.974	.580
50	15	.1	.210	.449	.143	.208	.212	.409	.182	.201	.146	.351	.091	.172
		.3	.808	.970	.821	.539	.800	.946	.827	.522	.697	.912	.697	.499
		.5	.999	1.000	.999	.800	.998	1.000	.999	.757	.999	1.000	.999	.784
100	30	.1	.288	.722	.189	.306	.336	.689	.260	.312	.211	.596	.119	.262
		.3	.974	1.000	.976	.781	.967	.999	.976	.764	.915	.995	.913	.732
		.5	1.000	1.000	1.000	.972	1.000	1.000	1.000	.969	1.000	1.000	1.000	.977
200	60	.1	.478	.950	.313	.494	.532	.949	.415	.463	.338	.869	.180	.410
		.3	.999	1.000	.999	.961	.999	1.000	1.000	.954	.995	1.000	.990	.944
		.5	1.000	1.000	1.000	.999	1.000	1.000	1.000	.999	1.000	1.000	1.000	.999
400	120	.1	.738	.998	.496	.716	.790	.998	.637	.695	.532	.991	.285	.590
		.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	1.000	1.000	1.000	.999
		.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

3.2. Minimum n for approximate normality of T_s and T_t

Both T_s and T_t are discrete statistics. When n is small we do not know their exact distributions. Thus it would be useful to study the minimum n that ensures good normal approximations of the exact distributions of T_s and T_t. The Shapiro-Francia W' test (Shapiro and Francia (1972)) is employed for testing departure from normality. The test assumes the form:

$$W' = \frac{(\sum_{i=1}^{N} b_i y_{(i)})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2},$$

where $\mathbf{b}' = (b_1, \ldots, b_N) = m'/(m'm)^{1/2}$ and $\mathbf{m}$ is the vector of expected values of standard normal order statistics. We employ W' to obtain the minimum value of n required for good normal approximations of the exact distributions of the weighted rank statistics. With $N = 100$ (100 points of T_s and T_t), and $p = 0.1, 0.2, 0.3$ and 0.5, we find that $n = 19, 9, 6$ and 6 are required for T_s, and $121, 18, 9$ and 6 are required for T_t.

Table 4. Powers of weighted Spearman's Rho, Kendall's Tau and top-down statistic with $p = 0.5$.

			N(0, 1)				Logistic(0, −1)				Extreme value dist.			
$n+1$	m	ρ	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T	r_w	τ_w	T_w	r_T
10	5	.1	.123	.109	.133	.092	.169	.103	.149	.091	.155	.098	.133	.084
		.3	.366	.255	.371	.196	.390	.240	.397	.194	.370	.212	.369	.188
		.5	.661	.478	.680	.333	.649	.467	.681	.330	.632	.418	.649	.331
		.7	.948	.970	.954	.551	.963	.967	.955	.538	.964	.959	.950	.545
20	10	.1	.183	.156	.145	.116	.188	.195	.173	.117	.156	.175	.139	.103
		.3	.576	.519	.537	.309	.589	.552	.566	.292	.517	.481	.484	.277
		.5	.912	.898	.899	.547	.911	.879	.896	.536	.893	.831	.866	.539
		.7	1.000	1.000	.999	.784	.999	1.000	1.000	.785	.999	1.000	.998	.816
30	15	.1	.207	.246	.168	.148	.244	.227	.209	.138	.200	.199	.152	.128
		.3	.739	.783	.689	.427	.728	.727	.693	.422	.654	.652	.605	.399
		.5	.983	.986	.974	.717	.974	.978	.968	.692	.962	.957	.945	.694
		.7	1.000	1.000	1.000	.930	1.000	1.000	1.000	.939	1.000	1.000	1.000	.957
50	25	.1	.286	.347	.227	.183	.318	.356	.253	.187	.250	.307	.188	.159
		.3	.897	.934	.838	.592	.891	.923	.847	.572	.820	.868	.744	.541
		.5	.999	1.000	.999	.891	.997	.999	.997	.864	.995	.997	.988	.860
		.7	1.000	1.000	1.000	.993	1.000	1.000	1.000	.994	1.000	1.000	1.000	.998
100	50	.1	.475	.600	.317	.284	.500	.589	.350	.284	.390	.507	.247	.244
		.3	.993	.999	.974	.850	.993	.999	.982	.831	.975	.995	.933	.805
		.5	1.000	1.000	1.000	.993	1.000	1.000	1.000	.989	1.000	1.000	1.000	.980
		.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	100	.1	.676	.842	.465	.450	.745	.864	.561	.443	.601	.789	.393	.377
		.3	1.000	1.000	.998	.983	1.000	1.000	.999	.981	.999	1.000	.995	.970
		.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	200	.1	.905	.985	.692	.685	.938	.985	.796	.684	.839	.961	.605	.585
		.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999
		.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

4. Application

In the following, we apply the rank tests to a data set from the identification of junior high gifted students (Kuo (1995)). The data shown in Figure 1 are from

111 7th grade students at Ho-Ping and Min-Sen Junior High Schools in Taipei, Taiwan in 1992.

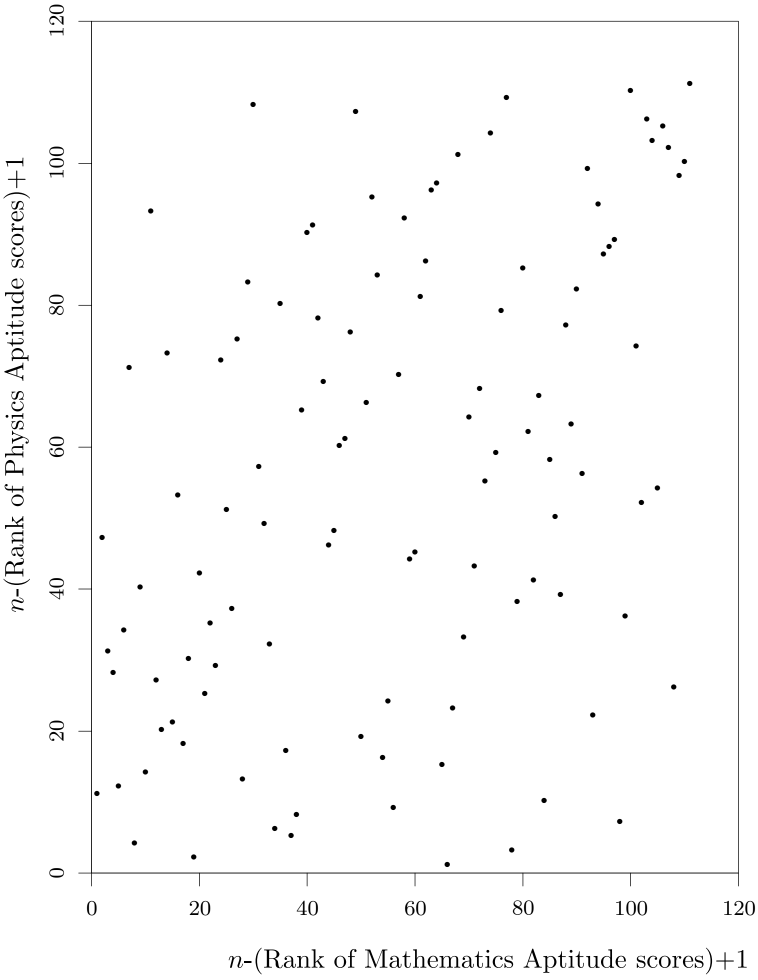

Figure 1. Mathematics scores versus Physics scores.

The paired ranks of Mathematics and Physics aptitude test scores are plotted against each other. According to Kuo (1995), the identification of the gifted is based on the results of multiple instruments, and what instruments to use is an important issue. Are Mathematics and Physics aptitude test scores correlated in the top ranks? Significant values of T_s and T_t imply that an aptitude test on one subject (say Mathematics) can be used as a screening instrument at the first stage, and then only those top-ranked in Mathematics take the Physics aptitude test at the next stage. The gifted may be chosen from those top-ranked in both (or a combination of) Mathematics and Physics tests, when non-significant values of T_s and T_t imply that both tests are needed for all students.

For $p = 0.1, 0.2, 0.3$ and 0.5, the values of T_s equal 15,997, 21,915, 21,220 and 22,518, respectively, and all of their P-values are less than 0.001. We can estimate the P-values by $P_{H_0}(Z \geq t_s/s.d.)$, where Z is a $N(0,1)$ r.v., t_s is the value of T_s and $s.d.$ is the asymptotic standard deviation of T_s (which can be obtained from Theorem 2). The estimated P-values are $.0008, < .0002$, .0004 and $< .0002$. To have small loss in the P-value and to save computation, we choose $p = 0.2$ for T_s. Similarly, for all four values of p, the values of T_t are 6.674, 9.375, 10.488 and 10.016, respectively. The estimated P-values are $.012, .006, .005$ and .006 respectively. This suggests the choice of $p = 0.2$. Both T_s and T_t reject H_0 at $\alpha = 0.05$ in this example. We conclude that the Physics aptitude test can be skipped by all students except those top-ranked in the Mathematics aptitude test.

5. Conclusion

The proposed rank tests are good for testing independence against a weighted contamination alternative. The general form of the LMPR test under the weighted contaminated alternative is derived. We show that the tests are asymptotic LMPR tests with respect to the Logistic and the Extreme Value families, respectively. The two statistics T_s and T_t generalize, respectively, Spearman's Rho and r_T in Iman and Conover (1987). Despite the fact that τ_w is not the LMPR test for either the Logistic or the Extreme Value family, it performs comparably to T_s and T_t in the simulation study. Thus Pitman and/or Bahadur asymptotic relative efficiencies (Pitman (1949) and Bahadur (1960)) of T_s, τ_w and T_t to r_T are of interest, but are deferred to a future study. Weighted concordance among b sets of rankings, $b > 2$, are often encountered in the real world, for instance when three or more techniques need to be compared in sensitivity analysis. Besides the top-down concordance measure studied by Iman and Conover (1987), other adequate measures require investigation. We leave these questions open.

Acknowledgements

The authors thank the Chair Editor, the Consulting Editor, an Associate Editor and an anonymous referee for careful readings and constructive suggestions that led to an improved presentation. This research was supported in part by NSC grant 87-2118-M001017 (Shieh).

Appendix

Proof of Theorem 1.

Recall that under H_1, we have

$$X_i = X_i^* + w(X_i^*)\Delta Z_i, \quad Y_i = Y_i^* + \Delta Z_i, \quad i = 1, \ldots, n,$$

where X_i^*, Y_i^* and Z_i are mutually independent, the densities of X^* and Y^* are known and denoted by f and g, respectively, and the distribution of Z_i is not specified. Here, we present a brief outline of the proof for the case $x = x^* + w(x^*)\Delta z$ is monotone in x^* and w is continuously differentiable. Let $x^* = s(x, \Delta z)$ be a solution of $x = x^* + w(x^*)\Delta z$. The Jacobian is

$$|J| = \left|\frac{\partial(x^*, y^*, z)}{\partial(x, y, z)}\right| = (1 + \Delta z w'(x^*))^{-1}.$$

Thus the joint p.d.f. of X and Y and their marginal distributions are given as

$$h_\Delta(x, y) = \int f(s(x, \Delta z))g(y - \Delta z)[1 + \Delta z w'(s(x, \Delta z))]^{-1} dM(z),$$

$$f_\Delta(x) = \int\int h_\Delta(x, y)dydM(z) = \int f(s(x, \Delta z))[1 + \Delta z w'(s(x, \Delta z))]^{-1} dM(z),$$

$$g_\Delta(y) = \int g(y - \Delta z)dM(z).$$

By symmetrization, we have

$$\begin{aligned}
&\frac{1}{\Delta^2}[h_\Delta(x, y) - f_\Delta(x)g_\Delta(y)] \\
&= \frac{1}{2\Delta^2}\Big\{\int\int \Big[f(s(x, \Delta z))[1 + \Delta z w'(s(x, \Delta z))]^{-1} - f(s(x, \Delta z')) \\
&\quad [1 + \Delta z' w'(s(x, \Delta z'))]^{-1}\Big][g(y - \Delta z) - g(y - \Delta z')]dM(z)dM(z')\Big\} \\
&= \frac{1}{2}\Big\{\int\int (z - z')^2 [1 + \Delta z w'(s(x, \Delta\xi))]^{-1} \Big[f'(s(x, \Delta\xi))w(s(x, \Delta\xi)) + \\
&\quad f(s(x, \Delta\xi))[1 + \Delta z w'(s(x, \Delta\xi))]^{-1} w'(s(x, \Delta\xi))\Big] g'(y - \Delta\eta)dM(z)dM(z')\Big\},
\end{aligned}$$

where both ξ and η fall in the interval (z', z). Thus

$$\lim_{\Delta\to 0} \frac{1}{\Delta^2}[h_\Delta(x, y) - f_\Delta(x)g_\Delta(y)] = \sigma_z^2 (wf)'(x)g'(y)$$

at each point (x, y) such that x and y are continuity points of $(wf)'(\cdot)$ and $g'(\cdot)$, respectively. Similar to page 77 of Hájek and Šidák (1967), we have that

$$\lim_{\Delta\to 0} \frac{1}{\Delta^2}\ [(n!)^2 Q_\Delta(\mathbf{R_x} = \mathbf{r}, \mathbf{R_y} = \mathbf{R}) - 1]$$

$$= \lim_{\Delta \to 0} (n!)^2 \sum_{k=1}^{n} \int \cdots \int_{\mathbf{R_x}=\mathbf{r}, \mathbf{R_y}=\mathbf{R}} \Big[(h_\Delta(x_k, y_k) - f_\Delta(x_k) g_\Delta(y_k))/(\Delta^2) \cdot \prod_{i=k+1}^{n} f_\Delta(x_i) g_\Delta(x_i) \prod_{j=1}^{k-1} h_\Delta(x_j, y_j) \Big] dx_1 \cdots dx_n dy_1 \cdots dy_n$$

$$= \sigma_z^2 \sum_{k=1}^{n} \Big[n! \int \cdots \int_{\mathbf{R_x}=\mathbf{r}} \Big([w(x_k) f(x_k)]'/f(x_k) \Big) \prod_{i=1}^{n} f(x_i) dx_1 \cdots dx_n \Big] \cdot \Big[n! \int \cdots \int_{\mathbf{R_y}=\mathbf{R}} \Big(g'(y_k)/g(y_k) \Big) \prod_{i=1}^{n} g(y_i) dy_1 \cdots dy_n \Big]$$

$$= \sigma_z^2 \sum_{k=1}^{n} \; a_n(r_k, w, f) b_n(R_k, g) \qquad \text{(A.1.1)}$$

$$\stackrel{\mathcal{DF}}{=} \sigma_z^2 \sum_{k=1}^{n} \; a_n(k, w, f) b_n(R_k, g),$$

where $\stackrel{\mathcal{DF}}{=}$ denotes equality in distribution. Further, the limit under the integral sign can be justified by the Dominated Convergence Theorem.

The following is a further justification of (A.1.1).

$$n! \int \cdots \int_{\mathbf{R_x}=\mathbf{r}} ((wf)'/f)(x_k) \prod_{i=1}^{n} f(x_i) dx_1 \cdots dx_n$$

$$= n! \int \cdots \int \Big([(wf)'/f](x_k) \; I_{[R(x_i)=r_i, i \neq k]} I_{[R(x_k)=r_k]} \Big) \prod_{i=1}^{n} f(x_i) dx_1 \cdots dx_n$$

$$= E\Big([(wf)'/f](x_{(r_k)}) \sum_{j=1}^{n} I_{[R(x_j)=r_k]} \Big) = E\Big([(wf)'/f](x_{(k)}) \Big).$$

Proof of Theorem 2.

$$\mathrm{Var}\,(T_s) = \mathrm{Var}\Big(\sum_{i=1}^{n} [w_i (i - \frac{n+1}{2})(R_i - \frac{n+1}{2})] \Big)$$

$$\stackrel{H_0}{=} E\Big[\sum_{i=1}^{n} w_i (i - \frac{n+1}{2})(R_i - \frac{n+1}{2}) \Big]^2$$

$$= \sum_{i=1}^{m} \Big[(i - \frac{n+1}{2})^2 \Big] \mathrm{Var}\,(R_1) + \Big[\sum_{1 \le i \neq j \le m} (i - \frac{n+1}{2})(j - \frac{n+1}{2}) \Big] \; \mathrm{Cov}\,(R_1, R_2) \qquad \text{(A.2.1)}$$

Now $w_i = I_{[i \le m]}$ and $m = [(n+1)p]$. Further, we have

$$E(R_1) = (\frac{n+1}{2}) \;\; , \;\; \mathrm{Var}\,(R_1) = (n^2 - 1)/12 \;\; ,$$

$$E(R_1R_2) = [(n+1)(3n+2)]/12 \quad \text{and} \quad \mathrm{Cov}\,(R_1, R_2) = -\frac{n+1}{12}.$$

Plugging these into (A.2.1), after some algebra we have

$$\mathrm{Var}\,(T_s) = [(n+1)m(3n^3 - 9mn^2 + 10m^2n - 3m^3 - n)]/144.$$

Thus

$$\mathrm{Var}\,(\sqrt{n}T_s/[n(n^2-1)/12]) \rightarrow p[3(1-p)^3 + p^2] \text{ as } n \rightarrow \infty.$$

Assuming that $I(f) < \infty$, $I(g) < \infty$, and applying Theorem V.1.6a in Hájek and Šidák (1967) to T_s, we obtain Theorem 2.

Proof of Theorem 3.

With $c_{in} = w_i(S_{n-i+1} - 1)$ and $a_n(R_i) = S_{n-R_i+1} - 1$, we can write $T_t = \sum_{i=1}^n c_{in}a_n(R_i)$. Since under H_0, $\{i\}$ and $\{R_i\}$ are independent, we may treat $T_t = \sum_{i=1}^n c_{in}a_n(R_i)$ as a linear rank statistic. Although c_{in} here depends on n and the c_i in Theorem V.1.6 of Hájek and Šidák (1967) does not, one can similarly prove that

$$T_t/\sigma_T \rightarrow N(0,1),$$

where $\sigma_T^2 = \mathrm{Var}\,(T_t)$.

Next, we proceed to compute the variance of T_t. Let $c = n - m + 1$. Under H_0,

$$\begin{aligned}
\mathrm{Var}\,(T_t) = E(T_t^2) &= \sum_{i=1}^n w_i^2(S_{n-i+1}-1)^2 E[(S_{R_1}-1)^2] \\
&\quad + \sum_{i \neq j} w_i w_j (S_{n-i+1}-1)(S_{n-j+1}-1)E[(S_{R_1}-1)(S_{R_2}-1)] \\
&= \sum_{i=c}^n (S_i - 1)^2 \{E(S_{R_1}-1)^2 - E[(S_{R_1}-1)(S_{R_2}-1)]\} \\
&\quad + [\sum_{i=c}^n (S_i-1)]^2 \; E[(S_{R_1}-1)(S_{R_2}-1)].
\end{aligned}$$

By straightforward calculation,

$$\sum_{i=c}^n \; (S_i - 1)^2 = \Big(\sum_{i=c}^n S_i^2 - 2\sum_{i=c}^n S_i + (n-c+1)\Big).$$

Note that

$$S_i \sim ln(n/i), \qquad \sum_{i=1}^n S_i = n, \qquad \sum_{i=1}^n S_i^2 = 2n - S_1, \qquad \text{(A.3.1)}$$

$$E[(S_{R_1}-1)^2]=\frac{1}{n}\sum_{i=1}^{n}S_i^2-1=1-S_1/n,$$

and

$$E[(S_{R_1}-1)(S_{R_2}-1)]=\Big\{[\sum_{i=1}^{n}(S_i-1)]^2-\sum_{i=1}^{n}(S_i-1)^2\Big\}/(n(n-1))=-(n-S_1)/(n^2-n). \tag{A.3.2}$$

By $S_i \sim ln(n/i)$, rectangular approximation and letting n tend to infinity, we have

$$(1/n)\sum_{i=c}^{n}S_i \to -\int_{1-p}^{1} ln\ y\ dy = p+(1-p)\ln(1-p), \tag{A.3.3}$$

and

$$(1/n)\sum_{i=c}^{n}S_i^2 \to -(1-p)(ln(1-p))^2-2\int_{1-p}^{1} ln\ y\ dy. \tag{A.3.4}$$

As $n \to\infty$, $m/n \to p$. This together with (A.3.1), (A.3.3) and (A.3.4) yield

$$(1/n)\sum_{i=c}^{n}(S_i-1)^2\Big(E[(S_{R_1}-1)^2]-E[(S_{R_1}-1)(S_{R_2}-1)]\Big)\to p-(1-p)ln^2(1-p). \tag{A.3.5}$$

Further, by (A.3.3) and (A.3.4), we have

$$A_n \equiv \Big[\sum_{i=c}^{n}(S_i-1)/n\Big]^2 \to (1-p)^2 ln^2(1-p), \tag{A.3.6}$$

as $n \to\infty$. Plugging (A.3.5), (A.3.2) and (A.3.6) into Var (T_t), we have Var $(T_t)/n \to p-(1-p)(2-p)ln^2(1-p)$ as $n \to\infty$. Hence Theorem 3 follows.

References

Bai, Z. D., Shieh, G. S. and Tsai, W. Y. (1997). Rank tests for independence — with a weighted correlation alternative. Technical Report No. C97-13, Institute of Statistical Science, Academia Sinica, Taiwan.

Bahadur, R. R. (1960). Stochastic comparison of tests. *Ann. Math. Statist.* **31**, 276-295.

Butler, A. K. (1986). Psychometric tools for knowledge engineering: correlational measures for evaluating structured selectors. Paper presented at AAA-I, Engineering Section, Philadelphia, August.

Galton, F. (1877). Typical laws of heredity. *Proc. Roy. Inst. Great Britain* **8**, 282-301.

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

Iman, R. L. (1987). Tables of the exact quantiles of the top-down correlation coefficient for $n = 3(1)14$. *Comm. Statist. Ser. A* **16**, 1513-1540.

Iman, R. L. and Conover, W. J. (1987). A measure of top-down correlation. *Technometrics* **29**, 351-57.

Kendall, D. G. (1938). A new measure of rank correlation. *Biometrika* **30**, 81-93.

Kuo, Ching-chih. (1995). The predictive validities of the instruments and the selective strategies used to identify the junior high gifted students. NSC report, Department of the Gifted Education, National Taiwan Normal University Taiwan.

Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. *Philos. Trans. Roy. Soc. London Ser. A* **187**, 253-318.

Pitman, E. J. G. (1949). *Lecture Notes on Nonparametric Statistical Inference*. Columbia University, New York.

Quade, D. and Salama, I. A. (1992). A survey of weighted rank correlation. In *Order Statistics and Nonparametrics: Theory and Applications* (Edited by P. K. Sen and I. A. Salama), 213-224. Elsevier Science Publishers B.V., Amsterdam.

Salama, I. A. and Quade, D. (1981). A nonparametric comparison of two multiple-regression prediction situations, Institute of Statistics, Mimeo Series # 1325, University of North Carolina, Chapel Hill.

Salama, I. A. and Quade, D. (1982). A nonparametric comparison of two multiple regressions by means of a weighted measure of correlation. *Comm. Statist. Ser. A* **11**, 1185-1195.

Savage, I. R. (1956). Contributions to the theory of rank order statistics— the two-sample case. *Ann. Math. Statist.* **27**, 590-615.

Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality. *J. Amer. Statist. Assoc.* **67**, 215-216.

Shieh, G. S. (1998). A weighted Kendall's tau statistic. *Statist. Probab. Lett.* **39**, 17-24.

Spearman, C. (1904). The proof and measurement of association between two things. *Amer. J. Psych.* **15**, 72-101.

Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan.

E-mail: gshieh@stat.sinica.edu.tw

Department of Mathematics, National University of Singapore, Singapore.

E-mail: matbaizd@leonis.nus.edu.sg

Division of Biostatistics, School of Public Health, Columbia University, U.S.A.

E-mail: wyt@biostat.columbia.edu

(Received May 1998; accepted August 1999)