Antonín Lukš

A nonparametric test of zero intrapair correlation

Aplikace matematiky, Vol. 28 (1983), No. 2, 91–102

Persistent URL: http://dml.cz/dmlcz/104009

A NONPARAMETRIC TEST OF ZERO INTRAPAIR CORRELATION

ANTONÍN LUKŠ

(Received November 5, 1980)

1. INTRODUCTION

Let

$$(1.1) \qquad (X_{11}, \ldots, X_{1n_1}), \ldots, (X_{k1}, \ldots, X_{kn_k})$$

be N $(= n_1 + \ldots + n_k)$ random observations with a continuous joint distribution. We have the model $X_{ij} = U_i + V_{ij}$; $i = 1, \ldots, k$; $j = 1, \ldots, n_i$, where U_i, V_{ij} are independent random variables, V_{ij} with a distribution function G. We shall test the null hypothesis H_0 that $U_1 = U_2 = \ldots = U_k = \Delta$ where Δ is a constant, *the hypothesis of independence*, against some of the following two alternatives:

(H_1) $U_i = \Delta_i$, $i = 1, \ldots, k$, where $\Delta_1, \Delta_2, \ldots, \Delta_k$

are constants, not all being equal. (*The hypothesis of difference in location* [3], p. 67).

(H_2) U_i, $i = 1, \ldots, k$, are random variables with a nondegenerate distribution function M. (*The hypothesis of dependence or heterogeneity*.) (Cf. [3], p. 75.)

H_0 may be interpreted as the hypothesis of no difference in location or that of homogeneity, respectively, according to if one tests H_0 against H_1 or H_2.

$H_0 \cup H_1$ and $H_0 \cup H_2$ are perhaps two probabilistic approaches to the one-way classification scheme. Both are very common, the former being known as the fixed effects model, the latter as the random effects model. To test H_0 against H_1 the Kruskal-Wallis test is widely used, along with the Wilcoxon two-sample test.

From the viewpoint of applications the fixed effects model is appropriate for small k([1], [2]), n_i large, and the random effects model useful for large k, n_i small.

If U_i, V_{ij} are normal variables, the corresponding submodel of $H_0 \cup H_1$ or of $H_0 \cup H_2$, respectively, is called a normal model.

The model $H_0 \cup H_1$ will not be treated in the sequel.

In the case of the model $H_0 \cup H_2$ we will apply the following measure of intraclass correlation, introduced by Rothery [4]:

$$(1.2) \qquad \varrho_c = P(X_{\beta l} < \min(X_{\alpha i}, X_{\alpha j}) \quad \text{or} \quad X_{\beta l} > \max(X_{\alpha i}, X_{\alpha j}))\,,$$

where $\alpha \neq \beta(\alpha, \beta = 1, \ldots, k)$, $i \neq j$ $(i, j = 1, \ldots, n_\alpha)$, $(l = 1, \ldots, n_\beta)$. This application is correct by virtue of the exchangeability of the $X_{\alpha i}$'s with i fixed in (1.1), i.e. the symmetry of their joint distribution. So each triple $(X_{\alpha i}, X_{\alpha j}, X_{\beta l})$ has the same density function.

Let

$$(1.3) \qquad (R_{11}, \ldots, R_{1n_1}), \ldots, (R_{k1}, \ldots, R_{kn_k})$$

denote the corresponding overall ranks of the pooled set of observations. Rothery [4] proposed the following estimate

$$(1.4) \qquad r_c = \sum_{\alpha=1}^{k} C_\alpha / S_3\,,$$

where

$$(1.5) \qquad C_\alpha = \tfrac{1}{2} n_\alpha (n_\alpha - 1)(N - n_\alpha) + \tfrac{1}{6} n_\alpha (n_\alpha^2 - 1) - \tfrac{1}{2} \sum_{i=1}^{n_\alpha} \sum_{j=1}^{n_\alpha} |R_{\alpha j} - R_{\alpha i}|\,,$$

$$(1.6) \qquad S_3 = \tfrac{1}{2} \sum_{\alpha=1}^{k} n_\alpha (n_\alpha - 1)(N - n_\alpha)$$

and denoted

$$(1.7) \qquad T = \sum_{\alpha=1}^{k} \sum_{i=1}^{n_\alpha} \sum_{j=1}^{n_\alpha} |R_{\alpha i} - R_{\alpha j}|\,.$$

He showed that r_c is an unbiased estimate of ϱ_c. He studied properties of the measure and its estimator for a normal model. He showed that the method provides a relatively powerful test of the null hypothesis in a normal population.

He considered an application when the observations are made on individuals chosen from k distinct families. Likewise, an application concerning twins coming from k distinct births (and families) had led me, simultaneously, to the derivation of a recurrent formula for calculation of the distribution of the rank statistic

$$d = \sum_{\alpha=1}^{k} |R_{\alpha 1} - R_{\alpha 2}|\,,$$

which equals $\frac{1}{2}T$ for $n_\alpha = 2(\alpha = 1, \ldots, k)$.

2. THE RANK STATISTIC

We shall now treat the case when $n_1 = n_2 = \ldots = n_k = 2$. Let us have k pairs of observations

$$(2.1) \qquad (x_{11}, x_{12}), \ldots, (x_{k1}, x_{k2}) .$$

Pooling these pairs into one set of data

$$(2.2) \qquad x_{11}, x_{12}, \ldots, x_{k1}, x_{k2} ,$$

we consider their ranks and regard them as a permutation

$$(2.3) \qquad R_{11}, R_{12}, \ldots, R_{k1}, R_{k2} ,$$

where R_{ij}, $i = 1, \ldots, k; j = 1, 2$, are the numbers $1, \ldots, 2k$.

We compute

$$(2.4) \qquad d = \sum_{i=1}^{k} |R_{i1} - R_{i2}| ,$$

or, which is the same,

$$(2.5) \qquad d = \sum_{i=1}^{k} |R_{i(1)} - R_{i(2)}| ,$$

or

$$(2.6) \qquad d = \sum_{i=1}^{k} |R_{(i1)} - R_{(i2)}| ,$$

where

$$(2.7) \qquad R_{1(1)}, R_{1(2)}, \ldots, R_{k(1)}, R_{k(2)}$$

stands for the permutation (2.3) ordered by necessary intra-pair transpositions so that

$$(2.8) \qquad R_{i(1)} < R_{i(2)} \quad \text{for all} \quad i = 1, \ldots, k$$

and

$$(2.9) \qquad R_{(11)}, R_{(12)}, \ldots, R_{(k1)}, R_{(k2)}$$

stands for the permutation (2.7) ordered by an extra-pair permutation so that, in addition to

$$(2.10) \qquad R_{(i1)} < R_{(i2)} \quad \text{for all} \quad i = 1, \ldots, k ,$$

we have also

$$(2.11) \qquad R_{(11)} < R_{(21)} < \ldots < R_{(k1)} .$$

3. DISTRIBUTION OF THE STATISTIC UNDER THE NULL HYPOTHESIS

As can be seen from (2.6), it will suffice to study only the random permutations (2.9) in what follows, i.e. the random permutations (2.9) of the numbers $1, \ldots, 2k$, satisfying (2.10) and (2.11).

It can be seen easily that there are

$$\frac{(2k)!}{k!\,2^k} = 1\,.\,3 \ldots (2k - 1) \tag{3.1}$$

possible permutations (2.9), each of them being equally probable under H_0.

The k-tuple of random variables $R_{(11)}, \ldots, R_{(k1)}$ will be called a lower (ordered) set, the k-tuple $R_{(12)}, \ldots, R_{(k2)}$ an upper set.

We derive easily that

$$d = \sum_{i=1}^{k} |R_{(i1)} - R_{(i2)}| = \sum_{i=1}^{k} (R_{(i2)} - R_{(i1)}) = \sum_{i=1}^{k} R_{(i2)} - \sum_{i=1}^{k} R_{(i1)} = \tag{3.2}$$

$$= 2 \sum_{i=1}^{k} [(2i - 1) - R_{(i1)}] + k\,,$$

i.e. that the lower set determines d. The formula (3.2) implies moreover that d takes on only the integral values of the same parity as k.

Let $R_{(11)}, R_{(21)}, \ldots, R_{(k1)}$ be a lower set. Then $R_{(k2)}$ belongs to the set $\{R_{(k1)} + 1, \ldots, 2k\}$, i.e. it can be chosen in $2k - R_{(k1)}$ ways.

The rank $R_{(k-1,2)}$ belongs to the set

$$\{R_{(k-1,1)} + 1, \ldots, 2k\} - \{R_{(k1)}, R_{(k2)}\}\,,$$

i.e. it can be chosen in $2k - R_{(k-1,1)} - 2 = 2(k - 1) - R_{(k-1,1)}$ ways.

The rank $R_{(k-2,2)}$ belongs to the set

$$\{R_{(k-2,1)} + 1, \ldots, 2k\} - \bigcup_{j=k-1}^{k} \{R_{(j1)}, R_{(j2)}\}\,,$$

i.e. it can be chosen in $2k - R_{(k-2,1)} - 4 = 2(k - 2) - R_{(k-2,1)}$ ways. Etc.

At last, $R_{(12)}$ belongs to the set

$$\{R_{(11)} + 1, \ldots, 2k\} - \bigcup_{j=2}^{k} \{R_{(j1)}, R_{(j2)}\}\,,$$

i.e. it can be chosen in $2k - R_{(11)} - 2(k - 1) = 2 - R_{(11)}$ ways. Now it is obvious that the number of variants for the upper set is given by the product

$$C = \prod_{i=1}^{k} (2i - R_{(i1)})\,, \tag{3.3}$$

as long as $R_{(i1)} < 2i$ for all $i = 1, \ldots, k$, and it equals 0 otherwise.

The number of the permutations (2.9) leading to a given d is given evidently by the formula

$$Q_k(d) = \sum \prod_{i=1}^{k} (2i - R_{(i1)}) \tag{3.4}$$

where the summation in $\sum$ extends over all

$$1 \leqq R_{(11)} < R_{(21)} < \dots < R_{(k1)} \leqq 2k$$

$$R_{(i1)} < 2i, \quad i = 1, \dots, k$$

$$k + 2\sum_{i=1}^{k}(2i - 1 - R_{(i1)}) = d\,.$$

More generally, denote

$$Q_k(d \mid \xi) = \sum \prod_{i=1}^{k} (2i - R_{(i1)}) \tag{3.5}$$

where the summation in $\sum$ extends over all

$$1 \leqq R_{(11)} < R_{(21)} < \dots < R_{(k1)} < \xi$$

$$R_{(i1)} < 2i\,, \quad i = 1, \dots, k$$

$$k + 2\sum_{i=1}^{k}(2i - 1 - R_{(i1)}) = d\,,$$

for ξ, $k + 1 \leqq \xi \leqq 2k + 1$, the number of the permutations (2.9) leading to the given d under the condition that $R_{(k1)} < \xi$, so that particularly $Q_k(d) = Q_k(d \mid 2k + 1)$.

It can be easily seen that

$$Q_k(d \mid \xi) = \sum_{R_{(k1)}=k}^{\xi-1} (2k - R_{(k1)})\, Q_{k-1}(d + 1 - 4k + 2R_{(k1)} \mid R_{(k1)})\,. \tag{3.6}$$

According to this definition

$$Q_1(d \mid 2) = 1 \quad \text{for} \quad d = 1\,, \quad Q_1(d \mid 2) = 0 \quad \text{for} \quad d \neq 1\,, \quad \text{integer.}$$

$$Q_1(d \mid 3) = 1 \quad \text{for} \quad d = 1\,, \quad Q_1(d \mid 3) = 0 \quad \text{for} \quad d \neq 1\,, \quad \text{integer.}$$

The formula (3.6) summarizes the following relations, which may be helpful for numerical computation.

$$\begin{aligned} Q_2(d \mid 3) &= 2Q_1(d - 3 \mid 2) \quad \text{for } d \text{ integer} \\ Q_2(d \mid 4) &= Q_2(d \mid 3) + 1\,.\,Q_1(d - 1 \mid 3) \quad \text{for } d \text{ integer} \\ Q_2(d \mid 5) &= Q_2(d \mid 4) \quad \text{for } d \text{ integer} \end{aligned} \tag{3.7}$$

$$\begin{aligned} Q_3(d \mid 4) &= 3Q_2(d - 5 \mid 3) \quad \text{for } d \text{ integer} \\ Q_3(d \mid 5) &= Q_3(d \mid 4) + 2Q_2(d - 3 \mid 4) \quad \text{for } d \text{ integer} \\ Q_3(d \mid 6) &= Q_3(d \mid 5) + 1\,.\,Q_2(d - 1 \mid 5) \quad \text{for } d \text{ integer} \\ Q_3(d \mid 7) &= Q_3(d \mid 6) \quad \text{for } d \text{ integer} \end{aligned} \tag{3.8}$$

and so on.

Further details of computation follow, after some practice, from Table 1, where certain characteristics of $Q_k(d \mid \xi)$ are indicated typographically. Observe, for example, a useful detail that $Q_k(d \mid \xi) = 0$ for $d > k^2$, and, particularly, that $Q_k(d) = 0$ for $d > k^2$, i.e. that the largest value taken on by the statistic d (e.g. for

$$R_{(11)} = 1\,, \quad R_{(12)} = 2k\,, \quad R_{(21)} = 2\,, \quad R_{(22)} = 2k - 1\,, \ldots, \quad R_{(k1)} = k\,, \quad R_{(k2)} = k + 1)$$

is k^2, which may be interesting not only for the computation.

In terms of (3.4) the probability distribution of d can be expressed as follows:

$$P(d \mid H_0) = \frac{Q_k(d)}{1\,.\,3\,.\,\ldots\,.\,(2k-1)}\,. \tag{3.9}$$

Table 1. Computation of numbers of permutations leading to a given value of the statistic d $Q_k(d \mid \xi)$

		$d=$															
$k=$	$\xi=$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	2	1															
	3	1															
	3		0		2												
2	4		1		2												
	5		1		2												
	4			0		0		0		6							
	5			0		2		4		6							
3	6			1		4		4		6							
	7			1		4		4		6							
	5				0		0		0		0		0		0		24
	6				0		0		0		6		12		18		24
4	7				0		2		8		14		24		18		24
	8				1		6		12		20		24		18		24
	9				1		6		12		20		24		18		24

4. TESTING THE HYPOTHESIS

It can be seen that under the alternative H_2 the statistic d tends to take on lower values.

The critical region

$$W_\alpha = \{d = \sum_{i=1}^{k} |R_{i1} - R_{i2}|\,;\quad d \leqq c(\alpha)\}\,, \tag{4.1}$$

where $c(\alpha)$ is, with respect to the fact that d is a discrete statistic, to be determined so that

$$(4.2) \qquad \begin{aligned} &P(d \leqq c(\alpha) \mid H_0) \leqq \alpha \\ &P(d \leqq c(\alpha) + 1 \mid H_0) > \alpha \end{aligned}$$

(See Figs. 1, 2.)

Values of $c(\alpha)$ are presented in Table 2 for $\alpha = 0{\cdot}1$, $0{\cdot}05$, $0{\cdot}025$, $0{\cdot}01$, $0{\cdot}005$ and $k = 5, \ldots, 20$.

Table 2. Percentage points $c(\alpha)$

k \ α	0·1	0·05	0·025	0·01	0·005
5	11	9	7	7	5
6	16	14	12	10	10
7	23	21	17	15	13
8	32	28	24	22	20
9	41	37	33	29	27
10	50	46	42	38	34
11	63	57	53	47	43
12	76	70	64	58	54
13	89	83	77	69	65
14	104	96	90	82	76
15	121	113	105	97	91
16	140	130	120	112	104
17	159	147	139	127	121
18	178	166	156	144	138
19	201	187	177	163	155
20	222	208	196	184	174

5. A MONTE CARLO POWER STUDY

The power of the test with critical region

$$\sum_{i=1}^{k} |R_{i1} - R_{i2}| \leqq c(\alpha)$$

should be compared with those of the tests with critical regions

$$\sum_{i=1}^{k} (R_{i1} - R_{i2})^2 \leqq c_1(\alpha) \quad \text{(the } \textit{Kruskal-Wallis} \text{ test)}\,,$$

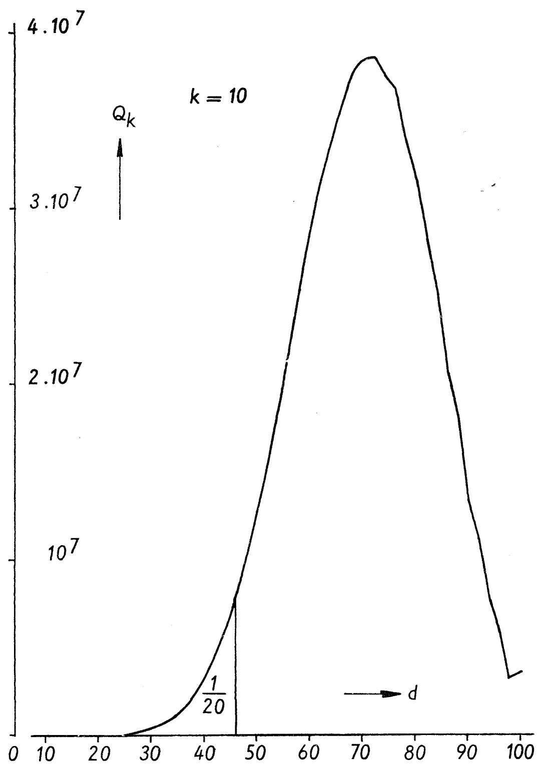

Fig. 1. Frequencies $Q_{10}(d)$. The statistic d takes on exclusively the even values from 10 to 100. The distribution of d for $k = 10$ is illustrated at the same time. It is approximately normal, especially on the left. The critical region of size 0·05 is indicated in the picture.

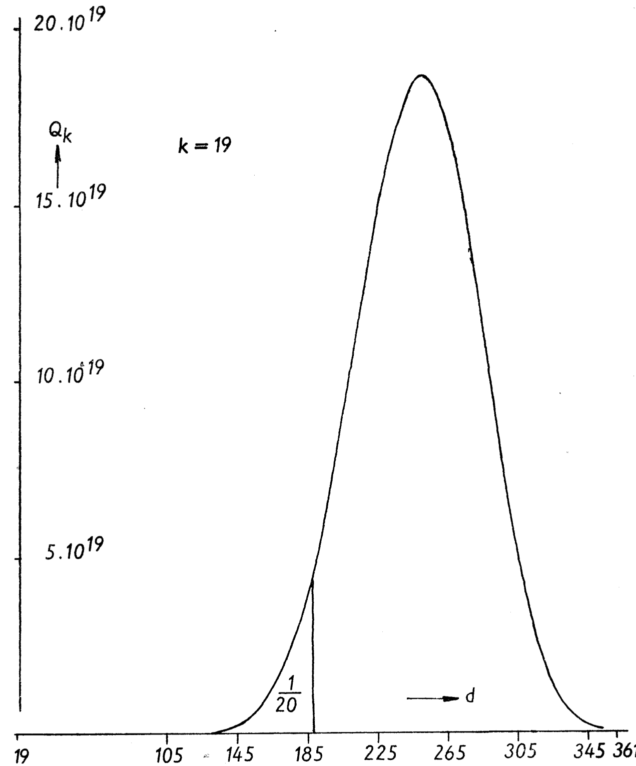

Fig. 2. Frequencies $Q_{19}(d)$. The statistic d takes on exclusively the odd values from 19 to $19^2 = 361$. The distribution of d for $k = 19$ is illustrated at the same time. It is fairly normal. The critical region of size 0·05 is indicated in the picture.

$$(\text{number of pairs separated by the sample median}) \leqq c_2(\alpha)$$
(the *median* test)

$$\sum_{i=1}^{k} \left(\Phi^{-1}\left(\frac{R_{i1}}{2k+1}\right) - \Phi^{-1}\left(\frac{R_{i2}}{2k+1}\right)\right)^2 \leqq c_3(\alpha)$$

(the *van der Waerden* test).

It is difficult to compare them in the interesting case of small samples because only different discrete sets of significance levels are available for each of the tests, although for large k one need not be excessively anxious about this fact.

A simulation study was carried out with the following aims and properties:

The power of the test with critical region $d \leqq c(\alpha)$ was compared with that of the Kruskal-Wallis test for $k = 5$ and 10. The normal model with ϱ equal to 0·1, 0·2, 0·4, 0·6 and 0·8 was used. Approximate percentage points $\hat{c}(0{\cdot}05)$ and $\hat{c}_1(0{\cdot}05)$ were calculated for the respective tests using 500 realizations of the model for each of the two values of k.

The power functions of the respective tests were estimated using 500 realizations of the model for different values of k and ϱ.

The results of this simulation study are presented in Table 3. For $k = 5$ there is no apparent difference in the power of the two tests, but for $k = 10$ the Kruskal-Wallis test seems to be more powerful.

Table 3. Estimated power functions

	$k = 5$		$k = 10$	
ϱ	$\hat{c}(0{\cdot}05) = 10$	$\hat{c}_1(0{\cdot}05) = 28$	$\hat{c}(0{\cdot}05) = 49$	$\hat{c}_1(0{\cdot}05) = 375$
0	0·034	0·044	0·044	0·050
0·1	0·034	0·038	0·076	0·090
0·2	0·052	0·062	0·130	0·144
0·4	0·090	0·108	0·270	0·294
0·6	0·208	0·202	0·544	0·596
0·8	0·442	0·438	0·898	0·918

6. CONSISTENCY OF THE r_c-TEST

Let us return to the general formulation as stated in the introduction. We shall prove the consistency of Rothery's r_c-test for H_0 against H_2 under some additional assumptions on the joint density function. The alternative is thus modified.

Let the following assumptions hold:
(a) G is absolutely continuous, $g := G'$ is continuous everywhere,
(b) g is positive everywhere,
(c) g is an even function,
(d) g is decreasing in $(0, +\infty)$.

Remark 1. $g(t)$ assumes the maximum in $(-\infty, +\infty)$ for $t = 0$ and is increasing for $t < 0$.

Lemma. *Let G be a distribution function with the properties* (a)–(d). *Then the function*

$$I(\Delta) = \int_{-\infty}^{\infty} (G^2(t - \Delta) + (1 - G(t - \Delta))^2)\, \mathrm{d}G(t)$$

of a real argument Δ (i) *has a derivative* I' *continuous everywhere,* (ii) *equals* 2/3 *for* $\Delta = 0$, (iii) $I(-\Delta) = I(\Delta)$ *for every* Δ, *and* (iv) $I' > 0$ *for* $\Delta > 0$.

Remark 2. Hence $I > 2/3$ for $\Delta \neq 0$, $I' = 0$ for $\Delta = 0$, and $I' < 0$ for $\Delta < 0$. Note that $I < 1$ for every Δ.

Proof. (i) Since

$$\frac{\partial}{\partial \Delta} (G^2(t - \Delta) + (1 - G(t - \Delta))^2) = (2 - 4G(t - \Delta))\, g(t - \Delta),$$

$$|(2 - 4G(t - \Delta))\, g(t - \Delta)| \leqq 2 \max g(t) = \text{const},$$

$$\int_{-\infty}^{\infty} 2 \max g(t) \,.\, g(t)\, \mathrm{d}t < +\infty,$$

$$I' = \frac{\mathrm{d}}{\mathrm{d}\Delta} \int_{-\infty}^{\infty} (G^2(t - \Delta) + (1 - G(t - \Delta))^2)\, g(t)\, \mathrm{d}t =$$

$$= \int_{-\infty}^{\infty} (2 - 4G(-\Delta))\, g(t - \Delta)\, g(t)\, \mathrm{d}t$$

is continuous everywhere.

(ii) The equality $I = 2/3$ for $\Delta = 0$ may be proved on substituting $u = G(t)$,

$$I = \int_0^1 (u^2 + (1 - u)^2)\, \mathrm{d}u .$$

(iii) Now,

$$I = \int_{-\infty}^{\infty} (G^2(t) + (1 - G(t))^2)\, g(t + \Delta)\, \mathrm{d}t .$$

On substituting $u = -t$ it follows that $I(-\Delta) = I(\Delta)$ for every Δ.

(iv) We have to prove that I' as expressed in (i) is positive for $\Delta > 0$. On substituting $t - \Delta = u$ we obtain

$$I' = \int_{-\infty}^{\infty} [2 - 4G(u)]\, g(u)\, g(\Delta + u)\, du ,$$

and now partition this integral into a sum of the two integrals $\int_{-\infty}^{0}$ and $\int_{0}^{\infty}$. In the second integral we substitute $u = -v$, thereafter we use in it the equalities $G(-v) = 1 - G(v)$, $g(-v) = g(v)$ following from the assumption (c) and this integral becomes

$$-\int_{-\infty}^{0} [2 - 4\, G(v)]\, g(v)\, g(\Delta - v)\, dv .$$

Replacing here v by u again and summing with the first integral, we obtain

$$I' = \int_{-\infty}^{0} [2 - 4\, G(u)]\, g(u)\, [g(\Delta + u) - g(\Delta - u)]\, du .$$

Due to the assumptions (c), (d), $\Delta > 0$ and since $u < 0$ in the domain of integration, it holds that $g(\Delta + u) > g(\Delta - u)$; so all the factors in the integral are positive and hence $I' > 0$. Q.E.D.

Remark 3. For every distribution function G it can be proved that

$$\lim_{\Delta \to \pm\infty} I(\Delta) = 1 .$$

Proof. *Viz.*,

$$|G^2(t - \Delta) + (1 - G(t - \Delta))^2| \leqq 1 = \text{const} ,$$

$$\int_{-\infty}^{\infty} 1 \,.\, dG(t) = 1 < +\infty ,$$

$$\lim_{\Delta \to \pm\infty} (G^2(t - \Delta) + (1 - G(t - \Delta))^2) = 1 .$$

Remark 4. The foregoing lemma may be applied to the normal, logistic, double exponential and Cauchy systems of distributions.

Theorem. *Let g be a density function with the properties* (a)–(d). *Then ϱ_c defined by* (1.2) *is greater than 2/3 under the modified alternative H_2.*

Proof. By the definition,

$$\varrho_c = P\{[(X_{\beta l} < X_{\alpha i}) \cap (X_{\beta l} < X_{\alpha j})] \cup [(X_{\beta l} > X_{\alpha i}) \cap (X_{\beta l} > X_{\alpha j})]\} =$$
$$= E_{X_{\beta l}} P\{[(X_{\beta l} < X_{\alpha i}) \cap (X_{\beta l} < X_{\alpha j})] \cup [(X_{\beta l} > X_{\alpha i}) \cap (X_{\beta l} > X_{\alpha j})]\, |X_{\beta l}\} =$$

$$= E_{X_{\beta l}} E_{U_\alpha} P\{[(V_{\alpha i} > X_{\beta l} - U_\alpha) \cap (V_{\alpha j} > X_{\beta l} - U_\alpha)] \cup$$
$$\cup [(V_{\alpha i} < X_{\beta l} - U_\alpha) \cap (V_{\alpha j} < X_{\beta l} - U_\alpha)] \mid X_{\beta l}, U_\alpha\} =$$
$$= E_{X_{\beta l}} E_{U_\alpha} \{G^2(X_{\beta l} - U_\alpha) + [1 - G(X_{\beta l} - U_\alpha)]^2\} .$$

Next,

$$\varrho_c = E_{V_{\beta l}} E_{U_\beta} E_{U_\alpha} \{G^2(V_{\beta l} + U_\beta - U_\alpha) + [1 - G(V_{\beta l} + U_\beta - U_\alpha)]^2\} =$$
$$= E_{U_\beta} E_{U_\alpha} \int_{-\infty}^{\infty} \{G^2(t + U_\beta - U_\alpha) + [1 - G(t + U_\beta - U_\alpha)]^2\} g(t)\,dt =$$
$$= E_{U_\beta} E_{U_\alpha} I(U_\alpha - U_\beta) .$$

It can be seen that the expected value of any function $I(\Delta)$ which is even, increasing for $\Delta > 0$ and has a derivative continuous everywhere, may equal $I(0)$ only if $P(\Delta = 0) = 1$, in our case only if $P(U_\alpha - U_\beta) = 1$, or in other words only if M is a degenerate distribution function. Otherwise $E\,I(\Delta) > I(0)$. The Theorem follows from the Lemma and Remark 2. Q.E.D.

Remark 5. The estimator r_c is consistent. Hence the consistency of the r_c-test follows.

Acknowledgments. The author would like to thank RNDr. *S. Komenda*, CSc. for the impetus, as well as Mrs *J. Jelínková*, who has prepared the drawings. He wishes to express his gratitude to RNDr. *Zbyněk Šidák*, DrSc. and a referee for their valuable comments on this paper and its draft.

References

[1] *J. Anděl:* Mathematical statistics. SNTL, Praha 1978.

[2] *G. Claus, H. Ebner:* Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen, Volk und Wissen. Berlin 1974.

[3] *J. Hájek, Z. Šidák:* Theory of rank tests. Academia, Praha 1967.

[4] *P. Rothery:* A nonparametric measure of intraclass correlation, Biometrika 66 (1979), 629–639.

Souhrn

NEPARAMETRICKÝ TEST NULOVÉ PÁROVÉ KORELACE

ANTONÍN LUKŠ

Autor aplikuje testové kriterium P. Rotheryho na statistickou analýzu pozitivní korelace symetrických dvojic pozorování. V tomto zvláštním případě dospívá k novým výsledkům. Práce končí obecným důkazem konzistence Rotheryho testu.

Author's address: RNDr. *Antonín Lukš*, Ústav pro výzkum vyšší nervové činnosti Univerzity Palackého, I. P. Pavlova 13, 775 20 Olomouc.