BULLETIN OF THE
GREEK MATHEMATICAL SOCIETY
VOLUME 28, 1987A

ANALOGIES BETWEEN RIEDWYL-TYPE AND LINEAR RANK TEST STATISTICS

BY

Charalambos Damianou

1. Introduction.

In the present paper we discuss the analogy betwen the two-sample linear rank tests and the one-sample Riedwyl-type tests. A two-sample linear rank statistic tests the null hypothesis that the two random samples have identical but otherwise arbitrary c.d.f. $F(x)$, whereas a Riedwyl-statistic tests the null hypothesis whether it is reasonable to approximate the data with a specified c.d.f. $F_0(x)$. A Riedwyl-type statistic is constructed analogously to a two-sample linear rank statistic where the second sample is replaced by some quantiles of $F_0(x)$.

Section 2 will be devoted to the necessary definitions and the basic notations used throughout this chapter. Section 3 will emphasize that the points of the support for the distribution of a Riedwyl-type statistic are the same as those for a rank test statistic.

Sections 4 and 5 will deal with the distribution and certain properties of the ranks in the Riedwyl-type situation and the linear rank-type situation, respectively. These two sections will show that the probabilities at the points of support for the two distributions are different. However, in Section 6, certain common characteristics for the distributions of the two types of test statistics will be obtained.

The paper concludes with a worked example which illustrates the similarities and the differences in the distributions for the two analogous test statistics.

2. Definitions and Notation

Let $X_1, X_2, \ldots, X_N$ be N independent random variables from a contintous distribution with c.d.f. $F(x)$; let $X_{(1)} < X_{(2)} < \ldots < X_{(N)}$ be the corresponding order statistics. The statistic R_i will be called the rank of X_i if $X_i = X_{(R_i)}$, provided that the R_i-th order statistic is uniquely defined. Assuming that $F(x)$ is continuous, the probability of obtaining tied observations (and thus tied order statistics) is zero. So in this case the

ranks $R_1, R_2, \ldots, R_N$ are well-difined; this assumption will be made throughout this paper.

A statistic $T = T(R_1, R_2, \ldots, R_N)$ which is a function of the original observations only through their ranks $R_1, R_2, \ldots, R_N$ is called a rank statistic. Such statistics form the basis of a substantial body of statistical inference; they provide alternatives to the classical statistics, especially for hypothesis testing.

An important sub-family of rank statistics are the so-called linear rank statistics.

Definition 2.1. Let $\{a(1), \ldots, a(N)\}$ and $\{c(1), c(2), \ldots, c(N)\}$, be two arbitrary vectors of constants, and let $(R_1, R_2, \ldots, R_N)$ be a vector of ranks. A statistic of the form

$$S = \sum_{i=1}^{N} c(i)\, a(R_i) \tag{2.1}$$

is called a linear rank statistic. The constants $a(1), \ldots, a(N)$ are called the scores and $c(1), \ldots, c(N)$ are termed the regression constants.

The main purpose of this paper is to study the linear rank statistics for testing two independent random samples, and to develop the analogy which arises if the second sample is replaced by quantiles of the specified hypothetical distribution for this sample. Thus for every two-sample linear rank test statistic there arises an "analogue" one-sample test statistic of the Riedwyl-type. We will examine whether the results known for linear rank tests can be adapted for the Riedwyl-type tests.

Case A: Linear rank tests

Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_{k-1}$ be two independent random samples from continuous distributions with c.d.f.'s $F(x)$ and $G(x)$, respectively. The null hypothesis of interest is $H_0^* : F(x) = G(x)$ for all x; it will be tested against either a one-tailed or a two-tailed alternative.

Let $N = n + k - 1$ denote the size of the combined sample and $Z_1^* < Z_2^* < \ldots < Z_N^*$ be the ordered pooled set of X and Y values. Since F and G are assumed to be continuous this ordering is certainly unique. Suppose that $R_1^*, R_2^*, \ldots, R_{k-1}^*$ denote the ranks of the Y's, and that $R_k^*, R_{k+1}^*, \ldots, R_N^*$ denote the ranks of the X's in the pooled set of Z^*'s, respectively. $\underset{\sim}{R}^*$ will stand for the vector of the ranks $(R_1^*, \ldots, R_{k-1}^*)$ of the Y's.

Putting

$$c(i) = 0 \quad \text{for} \quad i = k, \quad k + 1, \ldots, N, \tag{2.2}$$

in (2.1) gives the linear rank test statistics

$$S^* = \sum_{i=1}^{k-1} c(i)\, a(R_i^*), \tag{2.3}$$

which is a function solely of the ranks of the Y's.

Case B: Riedwyl-type tests

Suppose $G(x)$ is a specified continuous c.d.f., $F_0(x)$, completely known, i.e. $G(x) = F_0(x)$ for every x. Suppose the random sample $Y_1, Y_2, \ldots, Y_{k-1}$ is replaced by some quantiles $q_1, q_2, \ldots, q_{k-1}$ of $F_0(x)$ and $Z_1 < Z_2 < \ldots < Z_N$ be the pooled ordered "sample" of X's and q's. Note that this ordering is again certainly unique. Let $R_1, R_2, \ldots R_{k-1}$ be the ranks of the quantiles, and let $R_k, \ldots, R_N$ be the ranks of the X's in the combined set of Z's. We will use $\underset{\sim}{R}$ to denote the vector of ranks $(R_1, R_2, \ldots, R_{k-1})$. Substituting the same constants from (2.2) into (2.1) we obtain the test statistic

$$S = \sum_{i=1}^{k-1} c(i)\, a(R_i). \tag{2.4}$$

This is a one-sample Riedwyl-type statistic, analogous to S^*, for testing the null hypothesis $H_0 : F(x) = F_0(x) \ \forall\, x$.

In other words the variables and their obtained values with an asterisk will refer to the two-sample linear rank tests, in the sence defined above. The variables and their obtained values without an asterisk will refer to the one-sample Riedwyl-type tests.

Special cases:

If we let

$$c(i) = 1 \tag{2.5}$$

and

$$a(i) = \frac{n + k}{2} - i, \tag{2.6}$$

in (2.3) and (2.4), for $i = 1, 2, \ldots, k - 1$, then we obtain the test statistics

$$S^* = \frac{(n + k)(k - 1)}{2} - \sum_{i=1}^{k-1} R_i^* \tag{2.7}$$

and

$$S = \frac{(n + k)(k - 1)}{2} - \sum_{i=1}^{k-1} R_i, \tag{2.8}$$

respectively. The rank test statistic S^* is the Wilcoxon-Mann-Whitney (W-M-W) two-sample statistic (see Wilcoxon (1945) and Mann-Whitney (1947)), whereas the S is the Rey's 1/2 V(n, k) statistic, see Rey (1979). These two statistics will be compared in Section 6.

3. Computation of S^* and S statistics

The exact null distribution of any test statistic S^* (defined by (2.3)), or its analogue Riedwyl-type test statistic S (defined by (2.4)), may be determined from the distributions of the vector of the ranks $\underset{\sim}{R}^* = (R_1^*, R_2^*, \ldots, R_{k-1}^*)'$ and $R = (R_1, R_2, \ldots, R_{k-1})'$, respectively. The direct method of calculating the distribution of S^* or S consists in computing first their values s^* and s for all possible vectors $\underset{\sim}{r}^* = (r_1^*, r_2^*, \ldots, r_{k-1}^*)'$ and $\underset{\sim}{r} = (r_1, r_2, \ldots, r_{k-1})'$ of the rank values and then assigning to each resulting value the probability $P[\underset{\sim}{R}^* = \underset{\sim}{r}^*]$ or $P[\underset{\sim}{R} = \underset{\sim}{r}]$, respectively. If necessary, we sum up the probabilities for cases in which the resulting values s^* or s coincide.

For example, the discrete distribution of S^*, under H_0^*, is given by

$$P[S^* = s^*] = \Sigma s^* P[\underset{\sim}{R}^* = \underset{\sim}{r}^*], \tag{3.1}$$

where summation is over all the unordered subsets $(r_1^*, \ldots, r_{k-1}^*)$ of $k - 1$ integers taken (without replacement) from $\{1, 2, \ldots, N\}$ for which

$$\sum_{i=1}^{k-1} c(i)\, a(r_i^*) = s^*. \tag{3.2}$$

Similarly, the discrete distribution of S, under H_0, is obtained in the same manner. Hence (3.1) and (3.2) become

$$P[S = s] = \Sigma s P[\underset{\sim}{R} = \underset{\sim}{r}] \tag{3.3}$$

and

$$\sum_{i=1}^{k-1} c(i)\, a(r_i) = s, \tag{3.4}$$

respectively.

For the test statistic S^* there are $\binom{n+k-1}{k-1}$ possible arrangements of the x and y values, giving $\binom{n+k-1}{k-1}$ possible sets of vectors $\underset{\sim}{r}^* = (r_1^*, r_2^*, \ldots, r_{k-1}^*)'$ (see Lemma 5.1 below). Moreover, if the $Y_1, Y_2, \ldots, Y_{k-1}$ are replaced by quantiles then the above arrangements and sets of vectors are unaffected, i.e. there is one-to-one correspondence between the vectors $\underset{\sim}{r}^*$ and $\underset{\sim}{r}$ with values $r_i^* = r_i$ for all $i = 1, 2, \ldots, k-1$. Since formulae (3.2) and (3.4) involve the same functions of the ranks r_i^* and r_i, and the same regression constants, the points of the support of the distributions of S^* and S will be exactly the same.

Thus the points of support for any Riedwyl-type test statistics S may be obtained from the points of support for the corresponding rank statistic S^* and vice-versa.

4. Distribution of the rank vector $\underset{\sim}{R} = (R_1, \ldots, R_{k-1})'$

A Riedwyl-type test statistic is defined as a function of the deviations $d_i(n, k)$, $i = 1, 2, \ldots, k-1$, given by (4.1) below, for general $k \gtreqless n$. These deviations will be examined and their distribution theory will be obtained. A relationship between the deviations $d_i(n, k)$ and the ranks R_i, $i = 1, 2, \ldots, k-1$, will give the distribution and some properties of the ranks R_i.

4.1. The deviations d_i (n, k)

As a measure of the difference between the hypothetical c.d.f. $F_0(x)$ and the sample c.d.f. $F_n(x)$ Maag et al. (1973) considered the deviations

$$d_i(n, k) = F_0(q_1) - F_n(q_i), \quad i = 1, 2, \ldots, k-1 \tag{4.1}$$

where

$$q_i = F_0^{-1}\left(\frac{i}{k}\right), \quad i = 1, 2, \ldots, k-1, \tag{4.2}$$

are the $(k-1)$ quantiles of the hypothetical c.d.f. $F_0(x)$ (the c.d.f. which is implied by H_0). The q_i's divide the support of the r.v. X into k classes of equal probability. We will use F_0^{-1} to denote the inverse function of F_0. In order to avoid ties at the quantiles in (4.2) we must assume that the hypothetical c.d.f. $F_0(x)$ is continuous.

Assuming continuity for $F_0(x)$, the deviations $d_i(n, k)$ become

$$d_i(n, k) = i/k - F_n\{F_0^{-1}(i/k)\}, \quad i = 1, 2, \ldots, k-1 \tag{4.3}$$

For some purposes a more convenient formula than (4.3) is available.

Since $nF_n(x)$ is the number of X_i's less than or equal to x, it follows that

$$nF_n(q_i) = R_i - i, \quad i = 1, 2, \ldots, k-1 \tag{4.4}$$

where $q_1, q_2, \ldots, q_{k-1}$ are given by (4.2) and $R_1, R_2, \ldots, R_{k-1}$ are the ranks as defined in formula (2.4). Hence the deviations (4.4) become

$$d_i(n, k) = \frac{n+k}{nk} i - \frac{1}{n} R_i, \tag{4.5}$$

for $i = 1, 2, \ldots, k-1$.

The deviations $d_i(n, k)$ between an hypothetical c.d.f. $F_0(x)$ and an empirical c.d.f. $F_n(x)$ are illustrated in Figure 1 in the case $n = 10$ and $k = 6$.

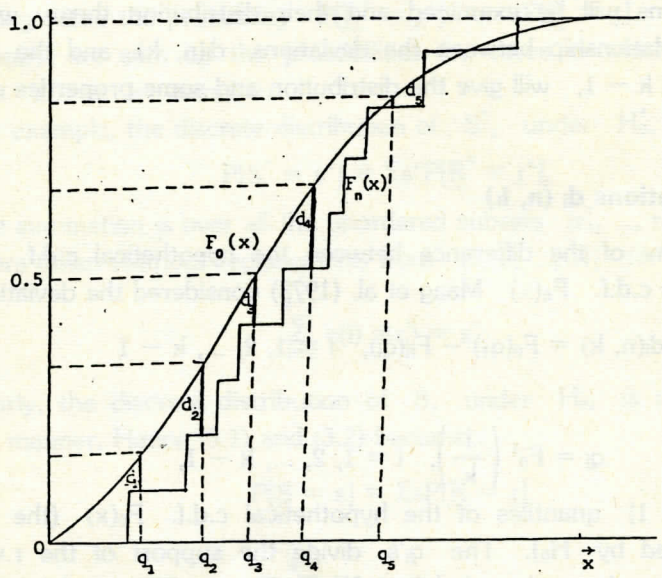

Figure 1. Deviations $d_1, d_2, \ldots, d_5$ between empirical $F_n(x)$ and hypothetical $F_0(x)$ distributions, when $n = 10$, $k = 6$.

4.2. General distribution theory for $d_i(n, k)$

To get the exact distribution of any Riedwyl-type test statistic we make use of the fact that the deviations $d_i(n, k)$ remain unchanged by any move of the sample point within the class (q_{i-1}, q_i). The probability that a sample point x falls in the class (q_{i-1}, q_i) is given by

$$p_i = P[X \in (q_{i-1}, q_i)]$$

$$= \int_{q_{i-1}}^{q_i} dF(x) = F(q_i) - F(q_{i-1})$$

i.e.,

$$p_i = F\left(F_0^{-1}\left(\frac{i}{k}\right)\right) - F\left(F_0^{-1}\left(\frac{i-1}{k}\right)\right), \quad i = 1, \ldots, k. \tag{4.6}$$

Therefore the problem reduces to the classical urn model where n balls are independently distributed into k urns, the probability of a ball falling into the i-th urn being p_i.

Let

$$D_i = nF_n\left(F_0^{-1}\left(\frac{i}{k}\right)\right), \quad i = 1, 2, \ldots, k \tag{4.7}$$

be the number of observations less than or equal to the quantile $q_i = F_0^{-1}\left(\frac{i}{k}\right)$. Then for fixed q_i, the r.v. D_i is binomially distributed with parameters n and $F\left(F_0^{-1}\left(\frac{i}{k}\right)\right)$. More generally, if we let

$$W_i = D_i - D_{i-1}, \quad i = 1, 2, \ldots, k \tag{4.8}$$

with $D_0 = 0$ and $D_k = n$, then the vector

$$\underset{\sim}{W} = (W_1, W_2, \ldots, W_k)'$$

has the multinomial distribution with n trials and probabilities p_i, $i = 1, 2, \ldots, k$ given by (4.6).

Consequently the vector

$$\underset{\sim}{W}_0 = (W_1, W_2, \ldots, W_{k-1})'$$

has the non-singular multinomial distribution,

$$R[W_1 = w_1, \ldots, W_{k-1} = w_{k-1}] = \frac{n!}{w_1!\, w_2! \ldots w_{k-1}!\, w_k!}\, p_1^{w_1} p_2^{w_2} \ldots p_k^{w_k}, \quad (4.9)$$

where

$$w_k = n - (w_1 + w_2 + \ldots + w_{k-1}) \quad \text{and} \quad w_i \geq 0, \quad i = 1, 2, \ldots, k$$

and the probabilities of p_i, $i = 1, \ldots, k$ are given by (4.6).

Recursively from (4.8) we find that

$$D_i = \sum_{t=1}^{i} W_t \quad \text{for} \quad i = 1, 2, \ldots k - 1 \quad (4.10)$$

Using (4.7) and (4.10), the deviations in (4.3) become

$$d_i(n, k) = \frac{i}{k} - \frac{1}{n} \sum_{t=1}^{i} W_t, \quad i = 1, 2, \ldots, k - 1 \quad (4.11)$$

In a matrix form (4.11) can be written as

$$\underset{\sim}{d} = \frac{1}{k}\, \underset{\sim}{e} - \frac{1}{n}\, L\, \underset{\sim}{W}_0 \quad (4.12)$$

where $\underset{\sim}{d}' = (d_i(n, k), \ldots, d_{k-1}(n, k))$, $\underset{\sim}{e}' = (e_1, \ldots, e_{k-1})$, $e_i = i$, and $L = (l_{ij})$ is a $(k - 1) \times (k - 1)$ matrix with elements

$$l_{ij} = \begin{cases} 1 & \text{for} \quad 1 \leq j \leq i \leq k - 1 \\ 0 & \text{otherwise} \end{cases}$$

From the multinomial distribution we obtain

$$E(W_i) = np_i, \qquad i = 1, 2, \ldots, k - 1 \quad (4.13)$$

$$\text{Var}(W_i) = np_i(1 - p_i), \quad i = 1, 2, \ldots, k - 1 \quad (4.14)$$

$$\text{Cov}(W_i, W_j) = -np_ip_j, \quad i \neq j \quad (4.15)$$

The corresponding results for the deviations $d_i(n, k)$ are obtained from (4.11). Thus,

$$E[d_i(n, k)] = \frac{i}{k} - E[D_i]/n$$

$$= \frac{i}{k} \sum_{t=1}^{i} p_t, \quad i = 1, 2, \ldots, k - 1 \quad (4.16)$$

and $\mathrm{Cov}[d_i(n, k), d_j(n, k)] = \frac{1}{n^2} \mathrm{Cov}(D_i, D_j)$

$$= \frac{1}{n}\left\{\sum_{l=1}^{i} p_l(1 - p_l) - \sum_{\substack{l=1}}^{i}\sum_{\substack{m=1 \\ m\neq l \\ 1\le i\le j\le k-1}}^{j} p_l p_m\right\} \tag{4.17}$$

where the p_i's are given by (4.6).

Under the null hypothesis, $H_0 : F(x) = F_0(x)$, the probabilities p_i from (4.6) become

$$p_i = \frac{1}{k}, \quad i = 1, 2, \ldots, k \tag{4.18}$$

Substituting (4.18) into (4.16) and (4.17) we obtain the expectation, the variance and the covariance under the null hypothesis for the deviations $d_i(n, k)$, $i = 1, 2, \ldots, k - 1$:

$$E[d_i(n, k)] = 0, \quad i = 1, 2, \ldots, k - 1 \tag{4.19}$$

$$\mathrm{Var}[d_i(n, k)] = \frac{i(k - i)}{nk^2}, \quad i = 1, 2, \ldots, k - 1 \tag{4.20}$$

and

$$\mathrm{Cov}[d_i(n, k), d_j(n, k)] = \frac{i(k - j)}{nk^2}, \quad 1 \le i \le j \le k - 1 \tag{4.21}$$

4.3. Distribution and properties of the ranks $R_1, R_2, \ldots, R_{k-1}$

Formula (4.5) gives the ranks R_i in terms of deviations $d_i(n, k)$, i.e.

$$R_i = \frac{n + k}{k} i - n d_i(n, k), \quad i = 1, 2, \ldots, k - 1 \tag{4.22}$$

In vector form, using (4.12),

$$\underset{\sim}{R} = L \underset{\sim}{W}_0 + \underset{\sim}{e} \tag{4.23}$$

with L, $\underset{\sim}{W}_0$ and $\underset{\sim}{e}$ as in (4.12).

Proposition 4.1. *The p.d.f. of the vector of ranks* $\underset{\sim}{R} = (R_1, R_2, \ldots, R_{k-1})'$ *is given by*

$$P[R_1 = r_1, \ldots, R_{k-1} = r_{k-1}] = \frac{n!}{w_1!\, w_2! \ldots w_k!} p_1^{w_1} p_2^{w_2} \ldots p_k^{w_k} \tag{4.24}$$

for $1 \leq r_1 < r_2 < \ldots < r_{k-1} \leq n + k - 1$, *and is zero, otherwise, where*

$$w_i = r_i - r_{i-1} - 1, \quad i = 1, 2, \ldots, k - 1; \; r_0 = 0$$
$$w_k = n - (w_1 + w_2 + \ldots + w_{k-1})$$
$$= n + k - 1 - r_{k-1}$$

and the probabilities p_i *are given by (4.6).*

Proof. Using (4.23) we see that the event

$$\{R_1 = r_1, R_2 = r_2, \ldots, R_{k-1} = r_{k-1}\}$$

implies

$$\{W_1 = r_1 - 1, W_2 = r_2 - r_1 - 1, \ldots, W_{k-1} = r_{k-1} - r_{k-2} - 1\}.$$

Since $W_i \geq 0$ for every $i = 1, 2, \ldots, k - 1$, (4.23) also gives (as expected) $R_1 < R_2 < \ldots < R_{k-1}$; the required result follows from (4.9).

Corollary 4.1. *Under the null hypothesis* $H_0 : F(x) = F_0(x)$ *the p.d.f. of* $\underset{\sim}{R}$ *is given by*

$$P[R_1 = r_1, \ldots, R_{k-1} = r_{k-1}] = \frac{n!}{w_1!\, w_2! \ldots w_k!\, k^n} \tag{4.25}$$

with the r_i*'s and* w_i*'s as defined above.*

Proof. The result follows immediately from Proposition 4.1 and formula (4.18).

Proposition 4.2. *Let* $R_1, R_2, \ldots, R_{k-1}$ *be the ranks given by (4.22). Under the null hypothesis* H_0.

(i)
$$E(R_i) = \frac{n + k}{k} i, \quad i = 1, 2, \ldots, k - 1 \tag{4.26}$$

(ii)
$$Var(R_i) = \frac{ni(k - i)}{k^2}, \quad i = 1, 2, \ldots, k - 1 \tag{4.27}$$

(iii)
$$Cov(R_i, R_j) = \frac{ni(k - j)}{k^2}, \quad 1 \leq i \leq j \leq k - 1 \tag{4.28}$$

and

(iv)
$$\rho(R_i, R_j) = \sqrt{\frac{i(k - j)}{j(k - i)}}, \quad 1 \leq i \leq j \leq k - 1. \tag{4.29}$$

Proof. The results (i) – (iii) are obtained directly from (4.22) and the results in (4.19) — (4.21), respectively. The correlation coefficient $\rho(R_i, R_j)$ is obtained from

$$\rho(R_i, R_j) = \frac{\mathrm{Cov}(R_i, R_j)}{\sqrt{\mathrm{Var}(R_i)}\sqrt{\mathrm{Var}(R_j)}}, \quad 1 \le i \le j \le k-1$$

and the results (ii) and (iii).

5. Distribution and properties of the ranks $R_1^*, \ldots, R_{k-1}^*$

Consider now the case A of Section 2 for the linear rank tests. Any two-sample rank test statistic, S^*, depends only on the set of values of $R_1^*, \ldots, R_{k-1}^*$ regardless of the order expressed by the indices. Therefore, to find the distribution of S^* we need to find only those arrangements of X's and Y's with resulting rank sets $\{R_1^*, \ldots, R_{k-1}^*\}$ in, say, increasing order, i.e. with $R_1^* < R_2^* < \ldots < R_{k-1}^*$.

This ordering is obtained if we assume, without loss of generality, that the Y_i's are given in increasing order, i.e. $Y_1 < Y_2 < \ldots < Y_{k-1}$. Equivalently to the definition of the ranks given in the beginning of Section 2, R_i^* (the rank of Y_i) is the number of observations in the combined set

$$\{X_i, Y_i \mid j = 1, 2, \ldots, n; \quad i = 1, 2, \ldots, k-1\} \tag{5.1}$$

less than or equal to Y_i, for $i = 1, 2, \ldots, k-1$.

In this case there is a similarity with the Riedwyl case, where the quantiles $q_1, q_2, \ldots, q_{k-1}$ are (by definition) ordered, i.e. $q_1 < q_2 < \ldots < q_{k-1}$.

In the present section we examine joint distributions and certain properties of the ranks $R_1^*, R_2^*, \ldots, R_{k-1}^*$ under the null hypothesis

$$H_0^* : F(x) = G(x) \quad \forall\, x.$$

Throughout this section the following assumptions are made.

Assumptions:

1) The $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_{k-1}$ are identically and independently distributed according to a common continuous distribution function F.

2) For the ranking of the observations in the set (5.1) we suppose that $Y_1 < Y_2 < \ldots Y_{k-1}$.

3) For $i = 1, 2, \ldots, k-1$, R_i^* denotes the rank of Y_i in the combined ranking of all the $N = n + k - 1$ observations. $\underset{\sim}{R}^*$ will denote the vector $(R_1^*, \ldots, R_{k-1}^*)'$ of ranks; it takes the values $r^* = (r_1^*, r_2^*, \ldots, r_{k-1}^*)$ where $r_1^* < r_2^* < \ldots < r_{k-1}^*$.

Lemma 5.1. *The number of ways of putting v alike objects into m different cells without any limitation is given by*

$$\binom{v+m-1}{v} = \binom{v+m-1}{m-1} \tag{5.2}$$

Proof. See e.g. Riordan (1958, p. 92).

Proposition 5.1. *The p.d.f. of the r.v. R_i^* under the assumptions stated at the beginning of this section, is given by*

$$p_i^j \equiv P[R_i^* = j] = \binom{j-1}{i-1}\binom{n+k-j-1}{k-i-1} \Big/ \binom{n+k-1}{n} \tag{5.3}$$

for $1 \leq i \leq k-1$, $i \leq j \leq n+i$, and is zero otherwise.

Proof. Suppose that the $(k-1)$ Y's are given first in their order $Y_1 < Y_2 < \ldots Y_{k-1}$. Then they define k spaces which may stand for k different cells. Let the n X's stand for the n stars (we suppose for the moment that all X's are alike). Then, according to Lemma 5.1 there are $\binom{n+k-1}{n}$ ways of putting the n stars into the k cells. After that we substitute each star with the $X_1, X_2, \ldots, X_n$, starting from the left. Therefore, we have $\binom{n+k-1}{n}$ arrangenents of X's and Y's from which we obtain $\binom{n+k-1}{n}$ possible $(k-1)$- tuples $(r_1^*, r_2^*, \ldots, r_{k-1}^*)$ for the rank vector $\underset{\sim}{R}^*$.

In order to have $R_i^* = j$, the Y_i observation must be in the j-th order in the ordered set (5.1). Also, $(i-1)$ Y's and $(j-i)$ X's must be smaller than Y_i, and $(k-i-1)$ Y's and $(n-j+i)$ X's must be larger, i.e. symbolically we have the following pattern:

$$Y_1\ Y_2\ \ldots\ Y_{i-1}\ Y_i\ Y_{i+1}\ \ldots\ Y_{k-1}$$
$$X_1, X_2, \ldots, X_{j-i} \uparrow X_{j-i+1}, \ldots, X_n$$
$$\text{j-th place}$$

The $(j - i)$ X's can be put in the i cells left of Y_i in $\binom{j-1}{i-1}$ ways; similarly, the $(n - j + i)$ X's right of Y_i can be put into the $(k - i)$ cells in $\binom{n+k-j-1}{k-i-1}$ ways. The total number of ways is given by their product, from which the proposition follows.

Corollary 5.1. *For the probability* $p_i^j = P[R_i^* = j]$ *of Proposition 5.1 to be a valid p.d.f. the following two conditions must hold:*

(i) $p_i^j \geq 0 \ \forall\, 1 \leq i \leq k-1;\ \ i \leq j \leq n+i,$

(ii) $\sum_{j=i}^{n+i} p_i^j = 1, \ \forall\, 1 \leq i \leq k-1.$

Proof. The first condition (i) is obvious. To prove (ii), it is sufficient to prove that

$$\sum_{j=i}^{n+i} \binom{j-1}{i-1} \binom{n+k-j-1}{k-i-1} = \binom{n+k-1}{n} \tag{5.4}$$

Putting $j - i = s$, the left-hand side (L.H.S.) of (5.4) becomes

$$\text{L.H.S.} = \sum_{s=0} \binom{s+i-1}{i-1} \binom{n+k-1-i-s}{k-i-1}$$
$$= \sum_{s=0} \binom{s+i-1}{s} \binom{n+k-i-1-s}{k-i-1}$$
$$= \binom{n+k-1}{n}$$

where we have used the hypergeometric summation formula

$$\sum_{s=0} \binom{p+s-1}{s} \binom{n-s}{m} = \binom{n+p}{m+p} \tag{5.5}$$

with $p = i$, $n = n + k - 1 - i$ and $m = k - i - 1$.

Corollary 5.2. *The probability that one of the r.v.'s* Y_i, $i = 1, 2, \ldots, k-1$, *in the set (5.1) has the rank* j, $i \leq j \leq n+i$ *is given by*

$$\sum_{i=1}^{k-1} P[R_i^* = j] = \frac{k-1}{n+k-1} \tag{5.6}$$

Proof. The event that there is a Y_i which has rank j implies that there is at least one R_i^* which takes the value j. Now {at least one $R_i^* = j$} implies

$$\{R_1^* = j\} \cup \{R_2^* = j\} \cup \ldots \cup \{R_{k-1}^* = j\}.$$

Because the events are mutually exclusive the required probability is $\sum_{i=1}^{k-1} P[R_i^* = j]$. Furthermore from (5.3)

$$\begin{aligned}\sum_{i=1}^{k-1} P[R_i^* = j] &= \binom{n+k-1}{n}^{-1} \sum_{i=1}^{k-1} \binom{j-1}{i-1} \binom{u+k-j-1}{k-i-1} \\ &= \binom{n+k-1}{n}^{-1} \sum_{s=0}^{k-2} \binom{r}{s} \binom{n+k-r-2}{k-2-s} \\ &= \binom{n+k-1}{n}^{-1} \binom{n+k-2}{k-2} = \frac{k-1}{n+k-1},\end{aligned}$$

where we have used for the last sum the summation formula

$$\sum_{s=0}^{k} \binom{r}{s} \binom{n-r}{k-s} = \binom{n}{k}.$$

Hence the result (5.6) follows.

Proposition 5.2. *The r-th ascending factorial moment of* R_i^*, $1 \leq i \leq k-1$, *is given by*

$$\mu_i^{[r]} = E[R_i^{*[r]} = i^{[r]}] = \binom{n+k+r-1}{n} \Big/ \binom{n+k-1}{n} \tag{5.7}$$

$$= i^{[r]}(n+k)^{[r]}/k^{[r]}$$

where $x^{[r]}$ *denotes the r-th ascending factorial, i.e.* $x^{[r]} = x(x+1) \ldots (x+r-1)$.

Proof. Using (5.3) it is sufficient to prove that

$$\sum_{j=i}^{k+i} j^{[r]} \binom{j-1}{i-1} \binom{n+k-j-1}{k-i-1} = i^{[r]} \binom{n+k+r-1}{n} \tag{5.8}$$

The L.H.S. of (5.8) becomes

$$\text{L.H.S.} = \sum_{j=1}^{n+i} i^{[r]} \binom{j+r-1}{i+r-1} \binom{n+k-j-1}{k-i-1}$$

$$= i^{[r]} \sum_{s=0}^{n} \binom{i+s+r-1}{s} \binom{n+k-i-s-1}{k-i-1}$$

which, from (5.5) with $p = i + r$, $m = k - i - 1$ and $n = n + k - i - 1$, becomes the R.H.S. of (5.8). The second expression in (5.7) (useful in applications) is easily obtained from the relation:

$$(x + y)!/y! = (y + 1)^{[x]}.$$

We prove next a proposition for the joint distribution function of any two rank variables R_i^* and R_j^*, for $1 \le i < j \le k - 1$. This proposition will be used in order to obtain the covariance of R_i^* and R_j^*.

Proposition 5.3. *The joint p.d.f. of two rank variables R_i^* and R_j^* under the assumptions stated at the beginning of this section is given by*

$$p_{ij}^{h\ell} \equiv P[R_i^* = h,\ R_j^* = \ell]$$

$$= \binom{h-1}{i-1} \binom{\ell-h-1}{j-i-1} \binom{n+k-\ell-1}{k-j-1} \Big/ \binom{n+k-1}{n} \tag{5.9}$$

for $1 \le i < j \le k-1$; $i \le h \le n+i$; $j \le \ell \le n+j$ and $h < \ell$, and is zero, otherwise.

$Y_1\ Y_2\ \ldots\ Y_{i-1}$ | $Y_i\ Y_{i+1}\ \ldots\ Y_{j-1}$ | $Y_j\ Y_{j+1}\ \ldots\ Y_{k-1}$

$X_1, X_2, \ldots, X_{h-i}$ | $X_{h-i+1}, \ldots, X_{\ell-j}$ | $X_{\ell-j+1}, \ldots, X_n$

h-th place

ℓ-th place

Group 1 Group 2 Group 3

Let v_i be the objects (X's) and m_i the cells (defined by the Y's) in group i for $i = 1, 2, 3$. Then we have

Group 1: $v_1 = h - i$, $m_1 = i$
Group 2: $v_2 = l - j - h + i$, $m_2 = j - i$
Group 3: $v_3 = n - l + j$, $m_3 = k - j$
Overall : $v = n$, $m = k$

Repeated use of Lemma 5.1 gives

$$P_{ij}^{hl} = \binom{v_1 + m_1 - 1}{v_1} \binom{v_2 + m_2 - 1}{v_2} \binom{v_3 + m_3 - 1}{v_3} \Big/ \binom{v + m - 1}{v}$$

from which the proposition follows immediately.

The procedures followed in Propositions 5.1 and 5.3 may be easily generalized to find the joint p.d.f. of any subset of the ranks $R_1^*, \ldots, R_{k-1}^*$.

The next proposition gives the p.d.f. of the vector $\underset{\sim}{R}^*$; i.e. the analogous result to that for the vector $\underset{\sim}{R}$, as given in Corollary 4.1 for the Riedwyl case.

Proposition 5.4. *The p.d.f. of the vector of ranks* $\underset{\sim}{R}^* = (R_1^*, R_2^*, \ldots, R_{k-1}^*)'$ *under the assumptions made earlier is given by*

$$P[R^* = r^*] = P[R_1^* = r_1^*, \ldots, R_{k-1}^* = r_{k-1}^*] = 1 \Big/ \binom{n + k - 1}{n} \qquad (5.10)$$

for $1 \leq r_1^* < r_2^* < \ldots < r_{k-1}^* \leq n + k - 1$, *and is zero otherwise.*

Proof. As in Proposition 5.3, we have now k groups with v_i the objects and m_i the cells in group i, $i = 1, 2, \ldots, k$. Overall we have again $v = n$ objects and $m = k$ cells. Perated use of Lemma 5.1 gives

$$P[R^* = r^*] = \prod_{i=1}^{k} \binom{m_i + v_i - 1}{v_i} \Big/ \binom{m + v - 1}{v}.$$

Since all $m_i = 1$, $i = 1, \ldots, k$, the result follows.

Proposition 5.5. *Under the null hypothesis* H_0^*, *the following results hold for the ranks* R_i^*, $i = 1, 2, \ldots, k - 1$:

(i) $$E[R_i^*] = \frac{(n + k)i}{k} \quad , \quad i = 1, 2, \ldots, k - 1 \qquad (5.11)$$

(ii) $$Var[R_i^*] = \frac{ni(k - i)(n + k)}{k^2(k + 1)}, \quad i = 1, 2, \ldots, k - 1 \qquad (5.12)$$

$$\text{(iii)} \quad Cov(R_i^*, R_j^*) = \frac{ni(k-j)(n+k)}{k^2(k+1)}, \quad 1 \le i \le j \le k-1 \tag{5.13}$$

$$\text{(iv)} \quad \rho(R_i^*, R_j^*) = \sqrt{\frac{i(k-j)}{j(k-i)}}, \quad 1 \le i \le j \le k-1 \tag{5.14}$$

Proof. The proof for (i) and (ii) follows immediately from Proposition 5.2 and the fact that

$$\mathrm{Var}(R_i^*) = \mu_i^{[2]} - \mu_i^{[1]} - (\mu_i^{[1]})^2.$$

For (iii) we have

$$\mathrm{Cov}(R_i^*, R_j^*) = E[R_i^* R_j^*] - E[R_i^*]\, E[R_j^*] \tag{5.15}$$

From Proposition 5.3

$$E[R_i^* R_j^*] = \sum_{\substack{h=i \\ n<l}}^{n+i} \sum_{l=j}^{n+j} h l p_{ij}^{hl}$$

$$= \binom{n+k-1}{n}^{-1} \sum_h \sum_l hl \binom{h-1}{i-1} \binom{l-h-1}{j-i-1} \binom{n+k-l-1}{k-j-1}$$

$$= i \binom{n+k-1}{n}^{-1} \sum_{l=j} l \binom{n+k-l-1}{k-j-1} \sum_{h=i} \binom{h}{i} \binom{l-h-1}{j-i-1}$$

$$= i \binom{n+k-1}{n}^{-1} \sum_{l=j} l \binom{n+k-l-1}{k-j-1} \sum_{s=0}^{n} \binom{i+s}{i} \binom{l-i-1-s}{j-i-1}$$

$$= i \binom{n+k-1}{n}^{-1} \sum_{l=j} l \binom{n+k-l-1}{k-j-1} \binom{l}{j}$$

$$= i \binom{n+k-1}{n}^{-1} \sum_{s=0}^{n} (s+j) \binom{s+j}{j} \binom{n+k-j-1-s}{k-j-1}$$

$$= ij \binom{n+k-1}{n}^{-1} \binom{n+k}{k} + i \binom{n+k-1}{n}^{-1} \sum_{s=0}^{n} s \binom{s+j}{j} \binom{n+k-j-1-s}{k-j-1}$$

$$= ij\,(n+k)/k + i \binom{n+k-1}{n}^{-1} \sum_{s=1}^{n} (j+1) \binom{s+j}{j+1} \binom{n+k-j-1-s}{k-j-1}$$

$$= ij\,(n+k)/k + i(j+1) \binom{n+k-1}{n}^{-1} \sum_{m=0}^{n-1} \binom{j+1+m}{j+1} \binom{n+k-j-2-m}{k-j-1}$$

$$= ij\,(n + k)/k + i(j + 1)\binom{n + k - 1}{n}^{-1}\binom{n + k}{k + 1}$$

$$= ij\,(n + k)/k + i(j + 1)\,n(n + k)/(k + 1)/k$$

from which

$$E[R_i^* R_j^*] = \frac{i(k + n)\{n(j + 1) + j(k + 1)\}}{k(k + 1)}. \tag{5.16}$$

Note that we have twise used the identity (5.5), and also have used the fact that

$$s\binom{s + j}{j} = (j + 1)\binom{s + j}{j + 1}.$$

Now, substituting (5.16) and (5.11) into (5.15) we obtain (5.13), i.e. the result (iii).

Finally, the proof of the last result (iv) is easily obtained from the results (ii) and (iii).

Remark 1. The formula (5.3) takes on an alternative form if we put $j = i + s$, for $s = 0, 1, \ldots, n$. Then

$$p_i^{i+s} \equiv P[R_i^* = i + s] = \binom{i + s - 1}{s}\binom{n + k - i - 1 - s}{n - s}\Big/\binom{n + k - 1}{n}$$

with $i = 1, 2, \ldots, k - 1$ and $s = 0, 1, 2, \ldots, n$.

This is the distribution of the number of exceedances studied by Gumbel and Schelling (1950).

Remark 2. Formulae (5.3) and (5.9), as well as the joint distribution of any subset of the ranks $R_1^*, R_2^*, \ldots, R_{k-1}^*$, can also be calculated from formula (5.10) by summation with respect to the other variables.

Remark 3. The p.d.f. of the rank R_i^* is symmetric in the following way:

$$P[R_i^* = j] = P[R_{k-1}^* = n + k - j];\quad i = 1, 2, \ldots, [k/2],\quad i \le j \le n + i.$$

Also, for the joint p.d.f. of R_i^* and R_j^* we have

$$P[R_i^* = h, R_j^* = l] = P[R_{k-i}^* = n + k - h,\quad R_{k-j}^* = n + k - l].$$

6. Results

Consider the test statistics S^* and S given by (2.3) and (2.4), respectively. In Section 3 we have seen that these statistics take the same values for every possible arrangement of the X's and Y's or the X's and q's, respectively, i.e. the vectors $\underset{\sim}{r}^*$ and $\underset{\sim}{r}$ have the same values. Thus, any algorithm which gives the points of the support for the one statistic may be used to give the points of support for the other.

Apart from this similarity, there are also other common properties for the two tests; these are obtained from the properties of the ranks R_i and R_i^*, $i = 1, 2, \ldots, k-1$ studied in Sections 4 and 5.

Formulae (4.25) and (5.10) show that the vectors $\underset{\sim}{R}$ and $\underset{\sim}{R}^*$ have different null distributions. The p.d.f. of $\underset{\sim}{R}$ is given by the multinomial distribution. Each $(k-1)$-tuple $(r_1, r_2, \ldots, r_{k-1})$ has a different probability. On the other hand, the p.d.f. of the vector $\underset{\sim}{R}^*$ is the discrete uniform distribution; every $(k-1)$-tuple $(r_1^*, \ldots, r_{k-1}^*)$ has the same probability, equal to $1 / \binom{n+k-1}{n}$. Therefore, the p.d.f. and the moments of the test statistic S^* are likely to be easier to obtain than the p.d.f. and moments of S.

Comparing the results of Propositions 4.2 and 5.5 we find that the rank variables R_i and R_i^* have the same expectations for every $i = 1, 2, \ldots, k-1$. However, the variance of R_i is always a constant multiple of the variance of R_i^* since

$$\mathrm{Var}(R_i) = \frac{k+1}{n+k}\,\mathrm{Var}(R_i^*) \quad \forall\, i = 1, 2, \ldots, k-1 \tag{6.1}$$

Note that $\mathrm{Var}(R_i) < \mathrm{Var}(R_i^*)$ since $\dfrac{k+1}{n+k} < 1$.

For every pair of rank variables (R_i, R_j) or (R_i^*, R_j^*) for $1 \le i \le j \le k-1$, the correlation coefficients are the same. But their covariances differ in the same way as the variances, i.e.

$$\mathrm{Cov}(R_i, R_j) = \frac{k+1}{n+k}\,\mathrm{Cov}(R_i^*, R_j^*), \quad 1 \le i \le j \le k-1 \tag{6.2}$$

These results give rise to some interesting results for the test statistics S and S^*. We distinguish the following cases:

a) $a(i) = ai + b$, $\forall i = 1, \ldots, k-1$ and a, b constants

The test statistics S and S^* have always the same means, i.e.

$$E(S) = E(S^*). \tag{6.3}$$

Their variance are related by

$$\mathrm{Var}(S) = \frac{k+1}{n+k}\,\mathrm{Var}(S^*).$$

Suppose, for example, that $c(i)$ and $a(i)$ are given by (2.5) and (2.6), respectively. Then S becomes the Rey's statistic $1/2V(n, k)$ (from (2.8)) and S^* becomes the W-M-W statistic W_s^* (from (2.7)). The first two moments of $V(n, k)$ can therefore be obtained from those of W_s^*. Since

$$E(W_s^*) = 0 \quad \text{and} \quad \mathrm{Var}(W_s^*) = \frac{n(k-1)(n+k)}{12} \tag{6.5}$$

it follows from (6.3) and (6.4) that

$$E[1/2V(n,k)] = 0 \quad \text{and} \quad \mathrm{Var}[1/2V(n,k)] = \frac{k+1}{n+k} \cdot \frac{n(k-1)(n+k)}{12},$$

and hence

$$E[V(n, k)] = 0 \quad \text{and} \quad \mathrm{Var}[V(n, k)] = \frac{n(k^2-1)}{3} \tag{6.6}$$

(see Rey (1979), p. 262).

b) $a(i) = i^2$ $\forall i = 1, 2, \ldots, k-1$

The mean value of S is obtained from that of S^* by the formula

$$E(S) = \frac{k+1}{n+k}\,E(S^*) + \frac{(n+k)(n-1)}{k^2}\sum_{i=1}^{k-1} i^2 c_i. \tag{6.7}$$

Suppose, for example, that $c(i) = 1$ and $a(i) = \{i - 1/2(n+k)\}^2$, $i = 1, 2, \ldots, k-1$. Then S^* becomes

$$M^* = \sum_{i=1}^{k-1} \{R_i^* - 1/2(n+k)\}^2, \tag{6.8}$$

i.e. the two-sample linear rank test proposed by Mood (1954) for testing

shifts in scale. The mean value of M^* was obtained by Mood and it is given by

$$E(M^*) = (k-1)(n+k)(n+k-2)/12. \tag{6.9}$$

An intuitive analogue of the Mood's test statistic of the Riedwyl-type is the statistic

$$M = \sum_{i=1}^{k-1} \{R_i - 1/2(n+k)\}^2. \tag{6.10}$$

The mean value of M can be obtained from that of M^* by the use of formula (6.7). After some algebra we find that

$$E(M) = \frac{k-1}{12k}\{(n+k)^2(k-2) + 2n(k+1)\}. \tag{6.11}$$

c) Let

$$a(i) = \frac{n+k}{2} - \left| i - \frac{n+k}{2} \right|, \quad i = 1, 2, \ldots, k-1$$

and regression constants $c(i) = 1$, $i = 1, 2, \ldots, k-1$. Then S^* becomes

$$S^* = \sum_{i=1}^{k-1} \left[\frac{n+k}{2} - \left| i - \frac{n+k}{2} \right| \right] \tag{6.12}$$

i.e. the rank test statistic for scale proposed by Ansari and Brandley (1960).

The analogous Riedwyl-type test statistic is

$$AB = \sum_{i=1}^{k-1} \left[\frac{n+k}{2} - \left| R_i - \frac{n+k}{2} \right| \right]$$

which might also be a useful test for scale.

7. Example

We close the paper with an example which illustrates the somewhat surprising results obtained in the previous sections.

Consider the case $n = 2$ and $k - 1 = 2$ (i.e. $k = 3$). Suppose we want to find the null distributions of the test statistics

$$S_1^* = \sum_{i=1}^{k-1} R_i^* \quad \text{and} \quad S_1 = \sum_{i=1}^{k-1} R_i$$

of the linear rank type and Riedwyl-type, respectively. There are

$$\binom{n+k-1}{n} = \binom{3}{2} = 6$$

arrangements to be examined in order to obtain the null distributions of these statistics. These arrangements, togehter with the associated probabilities, rank values and the associated values of the test statistics are given in Table 1.

Table 1.

Illustrated procedure to obtain the null distributions of the test statistics S_1^* and S_1 in the case $n = 2$, $k = 3$.

Arrangement	Linear rank case	Riedwyl case	Prob. in the linear rank case	Prob. in the Riedwyl case	Value of rank $R_1^* = R_1$	Value of rank $R_2^* = R_2$	Value of $S_1^* = S_1$
yyxx	$y_1y_2x_1x_2$	$q_1q_2x_1x_2$	1/6	1/9	1	2	3
yxyx	$y_1x_1y_2x_2$	$q_1x_1q_2x_2$	1/6	2/9	1	3	4
yxxy	$y_1x_1x_2y_2$	$q_1x_1x_2q_2$	1/6	1/9	1	4	5
xyyx	$x_1y_1y_2x_2$	$x_1q_1q_2x_2$	1/6	2/9	2	3	5
xyxy	$x_1y_1x_2y_2$	$x_1q_1x_2q_2$	1/6	2/9	2	4	6
xxyy	$x_1x_2y_1y_2$	$x_1x_2q_1q_2$	1/6	1/9	3	4	7

$S_1^* = S_1$	3	4	5	6	7
$P(S_1^*)$	1/6	1/6	2/6	1/6	1/6
$P(S_1)$	1/9	2/9	3/9	2/9	1/9

From Table 1 we find that

$$E(S_1^*) = 5, \quad Var(S_1^*) = 5/3$$

and

$$E(S_1) = 5, \quad Var(S_1) = 4/3$$

as we would expect from (6.3) and (6.4)

The joint p.d.f. of R_1^*, R_2^* and their marginals are given by

$R_1^* = i$ \ $R_2^* = j$	2	3	4	$\sum_j P[R_1^* = i,\ R_2^* = j]$
1	1/6	1/6	1/6	3/6
2	0	1/6	1/6	2/6
3	0	0	1/6	1/6
$\sum_j P[R_1 = i,\ R_2 = j]$	1/6	2/6	3/6	1

From this we find that

$$E(R_1^*) = 5/3, \qquad E(R_2^*) = 10/3, \qquad \mathrm{Var}(R_1^*) = 5/9,$$

$$\mathrm{Var}(R_2^*) = 5/9, \quad \mathrm{Cov}(R_1^*, R_2^*) = 5/18, \quad \rho(R_1^*, R_2^*) = 1/2$$

Note that results are in agreement with the results of Proposition 5.5. A similar table can be calculated for the joint p.d.f. of the ranks R_1, R_2 in the Riedwyl-case. The joint p.d.f. of R_1, R_2 and their marginals are given by

$R_1 = i$ \ $R_2 = j$	2	3	4	$\sum_j P[R_1 = i,\ R_2 = j]$
1	1/9	2/9	1/9	4/9
2	0	2/9	2/9	4/9
3	0	0	1/9	1/9
$\sum_i P[R_1 = i,\ R_2 = j]$	1/9	4/9	4/9	1

From this we find that

$$E(R_1) = 5/3, \qquad E(R_2) = 10/3, \qquad \mathrm{Var}(R_1) = 4/9,$$

$$\mathrm{Var}(R_2) = 4/9, \quad \mathrm{Cov}(R_1, R_2) = 2/9, \quad \rho(R_1, R_2) = 1/2.$$

These results check the findings of Proposition 4.2.

Comparisons of the results in the two cases check the findings of Section 6.

REFERENCES

1. Ansari, A.R. and Bradley, R.A. (1960). Rank-sum tests for dispersions. Ann. Math. Statist. 31, 1174-89.

2. Cumbel, E.J. and Von Schelling, H. (1950). The distribution of the number of exceedances. Ann. Math. Statist. 21, 247-62.

3. Maag, V.R., Streit, F. and Drouilly, P.A. (1973). Goodness-of-fit test for grouped data. JASA, 68, 462-65.

4. Mann, H.B. and Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Statist., 18, 50-60.

5. Mood, A.M. (1954). On the asymptotic efficiency of certain nonparametric two-sample tests. Ann. Math. Statist. 25, 514-22.

6. Rey, G. (1979). Properties and applications of the location test V(n, k). Biometrical Journal, 21, 259-76.

7. Riedwyl, H. (1967). Goodness-of-fit. JASA, 62, 390-98.

8. Riordan, J. (1958). An Introduction to Combinatorial Analysis. John Wiley, N.Y.

9. Wilcoxon, F. (1945). Individual comparison by ranking methods. Biometrics, 1, 80-83.

(Received by the editors, April 14, 1986)

Dr. Ch. Damianou Department of Mathematics, University of Athens, Athens, Greece.