CONTRIBUTIONS TO THE THEORY OF RANK ORDER STATISTICS—THE "TREND" CASE[1]

By I. Richard Savage[2]

Stanford University and *Center for Advanced Study in the Behavioral Sciences*

0. Introduction. In spirit this paper is a continuation of [5] and the techniques and terminology developed there will be used. Here we are concerned with the detailed relationships between the probabilities of rank orders under various "trend" hypotheses. The relationships found are of interest in themselves and in the theory of nonparametric tests of hypotheses.

Typically we shall be concerned with mutually independent random variables $X_1, \cdots, X_N$ such that X_i has a distribution function of the form $F(x - \theta_i)$ where the θ_i form an increasing sequence. Conditions are given under which one rank order is always more probable than another, one rank order is equally probable with another, and these results are translated into conditions for admissible rank order tests. References [1], [6], and [7] summarize information regarding large sample properties of nonparametric tests of this type of hypothesis.

In Section 1 two definitions of rank order are presented along with some "algebraic" properties of rank orders. Section 2 contains an enumeration of the hypotheses that we are concerned with. Section 3 presents theory and Section 4 contains applications.

1. Rank orders.

Definition: The rank order corresponding to the N distinct numbers $x_1, \cdots, x_N$ is the vector $r = (r_1, \cdots, r_N)$ where r_i is the number of x_j's $\leqq x_i$. r is a permutation of the first N integers. If in the definition the x's are replaced by random variables then R will be used instead of r. R will be defined with probability one when the underlying random variables have continuous distributions.

Definition: $r'L_{ij}r$ if $r'_k = r_k$ for $i \neq k \neq j$; $r'_i = r_j$, $r'_j = r_i$; and

$$(r_i - r_j)(i - j) > 0.$$

Thus, if $r = (2, 3, 6, 5, 4, 1)$ and $r' = (2, 5, 6, 3, 4, 1)$ then $r'L_{24}r$. We shall write $r'Lr$ as an abbreviation for $r'L_{ij}r$ or to denote that there is a chain of rank orders $r^1, \cdots, r^t, \cdots, r^T$ such that $r'L_{i_0j_0}r^1L_{i_1j_1} \cdots r^tL_{i_tj_t}r^{t+1} \cdots r^TL_{i_Tj_T}r$. Thus if $r = (2, 3, 6, 5, 4, 1)$ and $r' = (3, 5, 6, 4, 2, 1)$ then $r'Lr$, $T = 2$ and $r'L_{15}(2, 5, 6, 4, 3, 1)L_{45}(2, 5, 6, 3, 4, 1)L_{24}r$. For many of the hypotheses (see

Received November 8, 1956; revised July 18, 1957.

[1] This work was supported by an Office of Naval Research contract at Stanford University.

[2] Presently at the University of Minnesota.

Theorem 1) that we shall consider, $r'Lr$ will imply that rank order r' is *less* probable than rank order r.

DEFINITION: r^*Cr (rank order r^* is the complement of rank order r) if

$$r_i^* = N + 1 - r_{N+1-i} \qquad \text{for} \quad i = 1, \cdots, N.$$

If r^*Cr then rCr^*. The necessary and sufficient condition for rCr is that

$$r_i + r_{N+1-i} = N + 1.$$

If N^* is the largest integer less than or equal to $N/2$, then the number of self complementary rank orders is $(N^*)!2^{N^*}$ and the ratio of the number of rank orders to the number of self complementary rank orders is the product of the odd integers not greater than N. Thus most of the rank orders, for large N, occur in complementary pairs. Under a particular type of symmetry (see Theorem 5) complementary rank orders are *equally* probable.

Another definition of rank order can be given in the following manner. Let $\mathfrak{r} = (\mathfrak{r}_1, \cdots, \mathfrak{r}_N)$ where $\mathfrak{r}_i = j$ when the ith smallest of the numbers

$$(x_1, \cdots, x_N)$$

is x_j. The relationship between r and $\mathfrak{r}$ is: $r_a = b$ is equivalent to $\mathfrak{r}_b = a$ for $a, b = 1, \cdots, N$. It is easily verified that $r'Lr$ is equivalent to $\mathfrak{r}'L\mathfrak{r}$ and that r^*Cr is equivalent to $\mathfrak{r}^*C\mathfrak{r}$.

2. Hypotheses. Throughout we shall make the following assumption.

ASSUMPTION: The random variables $X_1, \cdots, X_N$ are mutually independent and each X_i has an absolutely continuous (w.r.t. Lebesgue measure) cumulative distribution function.

We shall let $F_i(x)$ denote the cumulative distribution function and $f(x, \theta_i)$ the density function of X_i. The θ_i's can be thought of as indices for the density functions but in many of the hypotheses they will correspond to parameters about which we shall make further assumptions.

H_0: There exists a cumulative distribution function $F(x)$ such that

$$F_i(x) \equiv F(x) \qquad \text{for} \quad i = 1, \cdots, N.$$

H_0 is our null hypothesis and under H_0 each rank order is equally probable.

H_1: The θ_i's are real valued and the following conditions hold:

1. $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_N$.
2. If $\theta_i < \theta_j$ and $x < y$, then

$$\begin{vmatrix} f(x, \theta_i) & f(x, \theta_j) \\ f(y, \theta_i) & f(y, \theta_j) \end{vmatrix} \geqq 0,$$

with strict $>$ for some $x < y$.

3. $f(x, \theta)$ is a continuous function in x for each θ.
4. The set of points on which $f(x, \theta)$ is positive does not depend on θ.

H_{2k} : The conditions of H_1 apply except $H_1(2)$ is replaced by: If

$$\theta_{i_1} < \theta_{i_2} < \cdots < \theta_{i_k}$$

and $x_1 < x_2 < \cdots < x_k$, then the determinant of $(a_{mn}) \geqq 0$, where $a_{mn} = f(x_m, \theta_{i_n})$. The inequality is strict for some $x_1 < x_2 < \cdots < x_k$.

H_3 : The θ_i's are real valued and the following conditions hold:
1. $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_N$.
2. $f(x, \theta_i) = g(\theta_i)h(x)e^{\theta_i x}$ where g and h are nonnegative functions.

H_4 : The θ's are real valued and the following conditions hold:
1. $0 < \theta_1 \leqq \cdots \leqq \theta_N$.
2. There is an absolutely continuous cumulative distribution function, $H(x)$, such that

$$F_i(x) = [H(x)]^{\theta_i}.$$

H_5 : The $\theta_i = i\theta > 0$, and $f(x, \theta_i) = f(x - i\theta) = f(i\theta - x)$.

The following relationships hold among H_1 to H_5. H_4 implies H_3 implies H_2 implies H_1. H_{22} is the same as H_1. H_5 is compatible with H_1, H_2, and H_3 but not H_4. The densities under H_1 satisfy the monotone likelihood ratio condition, the densities under H_2 have been described as being of Pólya-type k, the densities under H_3 are of exponential type [3], and the distributions under H_4 have been of interest in nonparametric inference ([4], [5]). H_5 implies that the density functions are symmetric about their medians. In practice, when a one-parameter family of distributions is assumed satisfactory for a particular problem, the family used often satisfies H_3. The Cauchy density with translation parameter satisfies only H_5. The distribution following Theorem 2k satisfies the conditions of H_{22} but not H_{23}. The extreme value distribution with log θ playing the role of the location parameter and the exponential distribution over the negative reals satisfy the assumptions of H_4.

A typical problem of interest is to form rank order tests of:

H_0—$X_1, \cdots, X_N$ are independently and identically distributed, against the alternative hypothesis

H_A—$X_1, \cdots, X_N$ are independently and normally distributed with common variance and $E(X_i) = a + y_i\theta$ where the y_i's are known and $\theta > 0$.

For hypotheses of the form H_A it is always possible to relabel the X's in order to make the y_i's a nondecreasing sequence. In doing the relabeling in order to preserve $\theta > 0$ it might also be necessary to replace all of the original observations by their negatives. Two sided alternatives would present no real difficulties. The removal of the assumption of continuous distribution functions and hence the possibility of ties would be more complicated.

3. Theoretical results. Due to the last remark made in Sec. 1 it will be apparent that Theorems 1, 2, and 5 and their corollaries are valid if the r definition of rank order is used instead of r. Theorems 3 and 4 and their corollaries are valid only in terms of r. The theorems, except Theorem 2k, with $k > 2$, give conditions for determining when the probability of one rank order is greater than the probability of another. These conditions yield necessary criteria for a rank order test to be admissible. Theorem 2k involves a relationship between $k!$ rank orders, which relationship, for $k > 2$, has not helped in determining admissibility of rank order tests.

THEOREM 1. *If* H_1, *then* $r'L_{ij}r$ *implies* $\Pr(R = r') < \Pr(R = r)$ *when*

$$\theta_i < \theta_j \quad \text{and} \quad r_i < r_j .$$

PROOF. A direct computation yields

$$\Pr(R = r) - \Pr(R = r') = \int_{-\infty < x_1 < \cdots < x_N < \infty} \cdots \int$$

$$\left\{ \left[\prod_{\substack{k=1 \\ i \neq k \neq j}}^{N} f(x_{r_k}, \theta_k) \right] \cdot [f(x_{r_i}, \theta_i) f(x_{r_j}, \theta_j) - f(x_{r_i}, \theta_j) f(x_{r_j}, \theta_i)] \left[\prod_{k=1}^{N} dx_k \right] \right\} .$$

The first bracket of the integrand is nonnegative since $f(x, \theta)$ is a density function. From assumption 2 of H_1 and since $\theta_i < \theta_j$ the second bracket of the integrand is always nonnegative and positive for some values of x_{r_i} and x_{r_j}, say u and v, such that $u < v$. From assumptions 3 and 4 of H_1 the whole integrand can be made positive in a region of the following type: x_{r_k} is near u for $r_k < r_j$ and x_{r_k} is near v for $r_k \geqq r_j$. Thus the integral is positive.

Without an assumption like H_1, Theorem 1 is false ([5], Sec. 5).

COROLLARY 1.1. *If* H_1, *then* $r'Lr$ *implies* $\Pr(R = r) > \Pr(R = r')$, *provided the* θ_{r_i} *corresponding to those* i *for which* $r_i \neq r_i'$ *are not all equal.*

COROLLARY 1.2. *In testing* H_0 *against* H_1 *with the added restriction*

$$\theta_1 < \theta_2 < \cdots < \theta_N$$

an admissible rank order test must have the following property: If $r'Lr$ *and the probability of rejecting* H_0 *is* >0 *when* $R = r'$, *then the probability of rejecting* H_0 *equals* 1 *when* $R = r$.

COROLLARY 1.3. *If* H_1 *and* $\theta_1 < \cdots < \theta_N$, *then* $\Pr(R_i = 1) > \Pr(R_{i+1} = 1)$ *for* $i = 1, 2, \cdots, N - 1$.

Let r^i be a typical rank order of the set of $k!$ possible rank orders which can be formed by permuting k integers in k positions of the vector describing a rank order. One element and the k positions determine a set. Denote such a set by $\mathcal{R}_k$. If the number of interchanges required to bring the k movable coordinates of r^i into increasing order is even let $c(r^i) = 1$ and otherwise let $c(r^i) = -1$. Thus for $N = 5$, $k = 3$, and positions 1, 4, 5 the following constitutes an example of an $\mathcal{R}_3$.

25314	$c = -1$
25341	$= 1$
15324	$= 1$
15342	$= -1$
45312	$= -1$
45321	$= 1$

THEOREM 2k: *If H_{2k}, then*

$$\sum_{r^i \varepsilon \mathfrak{a}_k} c(r^i) \Pr(R = r^i) \geqq 0,$$

with strict inequality when the k values of θ corresponding to the variable ranks are distinct.

PROOF. This theorem is proved in the same manner as Theorem 1.

To show that the results of Theorem 2k are not implied by the conditions of Theorem 1, consider the following density function

$$f(x, \theta) = \begin{cases} 0, & |x| > 1, \\ g(\theta)[100 + 10(x - \theta) - \epsilon^2(x - \theta)^2 - \epsilon^5(x - \theta)^3], & |x| \leqq 1, \end{cases}$$

where $|\theta| \leqq 1$, ϵ is a small fixed number and $g(\theta)$ is the normalization factor. For this density the conditions of Theorem 1 hold but the sign of the determinant in H_{23} is reversed. Thus not only is the condition of Theorem 2k not valid for this example but actually the inequality is reversed in the conclusion.

THEOREM 3. *If* (1) *the assumptions of H_3 hold,* (2) *rank orders $\mathfrak{r}''$ and $\mathfrak{r}$ are such that*

$$\sum_{j=1}^{i} \mathfrak{r}_j'' \geqq \sum_{j=1}^{i} \mathfrak{r}_j \qquad i = 1, \cdots, N$$

and the inequality is strict for at least one value of i and (3) *$\theta_i = i\theta$ where $\theta > 0$, then*

$$\Pr(\mathfrak{R} = \mathfrak{r}'') < \Pr(\mathfrak{R} = \mathfrak{r}).$$

PROOF. A direct computation yields

$$\Pr(\mathfrak{R} = \mathfrak{r}) - \Pr(\mathfrak{R} = \mathfrak{r}'') = \left[\prod_{i=1}^{N} g(\theta_i)\right] \int_{-\infty < x_1 < \cdots < x_N < \infty} \cdots \int \left[\prod_{i=1}^{N} h(x_i)\, dx_i\right] \cdot \left[\exp\left(\sum_{i=1}^{N} x_i \theta_{\mathfrak{r}_i}\right) - \exp\left(\sum_{i=1}^{N} x_i \theta_{\mathfrak{r}_i''}\right)\right].$$

It is sufficient to show that the last bracket is always positive or equivalently

$$\sum_{i=1}^{N} x_i(\theta_{\mathfrak{r}_i} - \theta_{\mathfrak{r}_i''}) > 0.$$

The identity

$$\sum_{i=1}^{N} x_i(\theta_{\mathfrak{r}_i} - \theta_{\mathfrak{r}_i''}) = \sum_{i=1}^{N-1} (x_i - x_{i+1}) \left[\sum_{j=1}^{i} (\theta_{\mathfrak{r}_j} - \theta_{\mathfrak{r}_j''}) \right]$$

yields the desired inequality, since:

(a) $$x_i - x_{i+1} < 0,$$

(b) $$\sum_{j=1}^{i} (\theta_{\mathfrak{r}_j} - \theta_{\mathfrak{r}_j''}) = \theta \sum_{j=1}^{i} (\mathfrak{r}_j - \mathfrak{r}_j'') \leqq 0,$$

and <0 for some i by assumptions 2 and 3.

Assumption 2 of Theorem 3 is equivalent to: $\sum_{j=i}^{N} \mathfrak{r}_j \geqq \sum_{j=i}^{N} \mathfrak{r}_j''$ for $i = 1, 2, \cdots N$, and the inequality is strict for some i. Incidentally, assumption 2 of Theorem 3 does not imply $\mathfrak{r}''L\mathfrak{r}$. This can be seen by examining

$$\mathfrak{r} = (2, 5, 1, 3, 4) \quad \text{and} \quad \mathfrak{r}'' = (3, 4, 2, 5, 1).$$

On the other hand, $\mathfrak{r}'L\mathfrak{r}$ does imply assumption 2 of Theorem 3 in the obvious direction.

COROLLARY 3.1. *In testing H_0 against H_3 (with the added restriction $\theta_i = i\theta > 0$) an admissible rank order test must be such that if rank orders $\mathfrak{r}''$ and $\mathfrak{r}$ satisfy assumption 2 of Theorem 3, then if the probability of rejecting H_0 is positive when $\mathfrak{r}''$ occurs, the probability of rejecting H_0 when $\mathfrak{r}$ occurs must be one.*

THEOREM 4. *Under H_4,*

$$\Pr(\mathfrak{R} = \mathfrak{r}) = \left(\prod_{i=1}^{N} \theta_i\right)\left[\prod_{i=1}^{N}\left(\sum_{j=1}^{i} \theta_{\mathfrak{r}_j}\right)\right]^{-1}.$$

PROOF. See Theorem 7a.1 and proof in [5].

COROLLARY 4.1. *In testing H_0 against H_4 with the added restriction $\theta_i = i\theta$ the uniformly most powerful rank order test is based on large values of the statistic*

$$T_3(r) = \prod_{i=1}^{N}\left(\sum_{j=1}^{i} \mathfrak{r}_j\right)^{-1}.$$

In Corollary 4.1 "uniformly" refers both to θ and to $H(x)$.

THEOREM 5. *Under H_5, if r^*Cr then* $\Pr(R = r^*) = \Pr(R = r)$.

PROOF.

$$\Pr(R = r) = \int_{-\infty < x_1 < \cdots < x_N < \infty} \cdots \int \prod_{i=1}^{N} [f(x_{r_i} - i\theta)\, dx_i].$$

Now make the change in variables $x_i = (N + 1)\theta - y_{N-i+1}$ and obtain

$$\Pr(R = r) = \int_{-\infty < y_1 < \cdots < y_N < \infty} \cdots \int \prod_{i=1}^{N} [f(-y_{N-r_i+1} + \theta(N + 1 - i))\, dy_i]$$

$$= \int_{-\infty < y_1 < \cdots < y_N < \infty} \cdots \int \prod_{i=1}^{N} [f(y_{r_i^*} - i\theta)\, dy_i] = \Pr(R = r^*).$$

TABLE 1

Admissibility properties of rank order tests of trend

1. Hypothesis	H_1	H_3	H_5
2. Theorem	Cor. 1.2	Cor. 3.1	Th. 5
3. Condition	$r'Lr$[3,4]	$\sum_{j=1}^{i} (\mathfrak{r}_j'' - \mathfrak{r}_j) \geqq 0$[5,6]	r^*Cr[7,8]
Statistic[1,2]			
$T_1(r) = \sum_{i=1}^{N} i\mathfrak{r}_i$	+	+	+ [6][9]
$T_2(r) = \sum_{i=1}^{N} \mathfrak{r}_i E_{iN}$	+	+	+
$T_3(r) = \prod_{i=1}^{N} \left[\sum_{j=1}^{i} \mathfrak{r}_j \right]^{-1}$	+	+	−
$T_4(r) = \sum_{i=1}^{N} \sum_{j=1}^{N} d(r_i, r_j)$	+	−	+ [6]
$T_5(r) = \sum_{i=1}^{N-1} d(r_i, r_{i+1})$	−	−	+ [6]
$T_6(r) = \sum_{\substack{i=n+1 \\ (N=2n)}}^{N} d(n, r_i)$	0	−	+ [1]
$T_7(r) = \sum_{\substack{i=1 \\ (N=2n)}}^{n} (N - 2i + 1) d(r_i, r_{N-i+1})$	−	−	+ [1]
$T_8(r) = \sum_{\substack{i=1 \\ (N=3n)}}^{n} d(r_i, r_{2n+i})$	−	−	+ [1]
$T_9(r) = \sum_{\substack{i=1 \\ (N=2n)}}^{n} d(r_i, r_{n+i})$	−	−	+ [1]
$T_{10}(r) = \sum_{i=1}^{N-1} [d(\max_{1 \leq j \leq i} r_j, r_{i+1}) - d(r_{i+1}, \min_{1 \leq j \leq i} r_j)]$	−	−	− [2]
$T_{11}(r) = T_{10}(r) - T_{10}(r^*), (r_i^* = r_{N-i+1})$	−	−	+ [2]

[1] For each statistic large values are critical for the alternatives of Sec. 2.

[2] Define E_{iN} as the expected value of the ith smallest observation in a sample of N from a normal distribution with mean 0 and variance 1. Also define $d(x, y)$ as 1 if $x < y$ and as 0 if $x \geqq y$.

[3] If $r'Lr$ implies $T(r) > T(r')$ the symbol + is recorded. If $r'Lr$ implies $T(r) \geqq T(r')$ the symbol 0 is recorded. If there exists r' and r such that $r'Lr$ and $T(r) < T(r')$ the symbol − is recorded.

[4] The results are easily obtained. The positive results are found by first examining the

In the notation of [5] for the two-sample case define z^*Cz to mean that

$$z_i^* = 1 - z_{N-i+1}.$$

Then, under the assumptions of symmetry and translation, z^*Cz implies

$$\Pr(Z = z^*) = \Pr(Z = z).$$

The proof is much like that of Theorem 5. Theorem 6.1 of [5], for the two-sample case, is implied by Theorem 1.

4. Applications. Many rank order tests have been proposed for the hypotheses of Sec. 2. At the present time we present a catalogue, far from complete, of such tests. Also included are some new tests. For all the tests listed large values of the test statistic are critical for the alternatives under consideration. Information regarding these tests is summarized in Table 1.

The statistic T_3 was introduced in Corollary 4.1. The statistic T_2 yields the rank order test whose power function has the largest derivative at $\theta = 0$ in testing H_0 against the alternative that $X_1, \cdots, X_N$ are independently distributed each with a normal distribution having common variance and

$$E(X_i) = a + i\theta, \qquad \theta > 0.$$

The symbol + in the column marked $r'Lr$ means that the test statistic satisfies the admissibility condition of Corollary 1.2 and the symbol − means that the test fails to satisfy this condition. The symbol 0 means that $T(r') = T(r)$ can occur when $r'Lr$ and thus Corollary 1.2 can be useful in discriminating between tied values of T. Positive results are obtained for those test statistics which make intercomparisons between all of the coordinates of the rank orders and negative results correspond to those test statistics whose structure is rather simple and not all of the intercomparisons are made.

The symbol + in the column labeled $\sum (\mathfrak{r}_j'' - \mathfrak{r}_j) \geqq 0$ means that the corresponding test statistic satisfies the admissibility condition of Corollary 3.1 and the symbol − signifies the statistic does not satisfy this condition. The results are like those for the $r'Lr$ column with a few more negative results since the con-

---

L_{ij} relationship. The negative results follow from counter examples. Thus for T_5 consider $r = (3, 2, 5, 4, 1)$ and $r' = (5, 2, 3, 4, 1)$.

[5] If $\sum_{j=1}^{i} (\mathfrak{r}_j'' - \mathfrak{r}_j) \geqq 0$ for all i and strict inequality for some i implies $T(r) > T(r'')$ the symbol + is recorded and the symbol − is scored if for some r'' and r satisfying the partial sums condition $T(r'') > T(r)$.

[6] The positive results for T_1 and T_2 are obtained in the same manner as the proof of Theorem 3. The positive result for T_3 is implied by Corollary 4.1. The negative results are obtained by constructing counter examples. Thus for T_4 consider $\mathfrak{r} = (1, 8, 2, 7, 6, 5, 4, 3)$ and $\mathfrak{r}'' = (4, 5, 3, 6, 7, 8, 1, 2)$.

[7] The symbol + is recorded if r^*Cr implies $T(r) = T(r^*)$ and the symbol − is recorded if for some r^* and r we have r^*Cr and $T(r) \neq T(r^*)$.

[8] These results are trivial.

[9] References are given to places where the test statistic has been used for the types of alternatives considered in Sec. 2 without attempting to reflect priority of publication. Reference [6] summarizes large sample efficiencies for many of the tests.

dition to be filled is stronger. The most interesting of these results is the − for T_4 which is essentially Kendall's tau. Thus for some levels of significance Kendall's tau is inadmissible amongst the class of rank order tests when considering trend in exponential alternatives.

The symbol + in the column labeled r^*Cr means that r^*Cr implies $T(r^*) = T(r)$. Under the conditions of Theorem 5 which frequently hold in practice this is a reasonable condition, i.e., rank orders which are equally probable give the same value for the test statistic. The sole negative result corresponds to T_3 which is the optimum statistic for a class of alternatives not included in the alternatives considered in Theorem 5.

Since the two forms of rank order, r and $\mathfrak{r}$, are equivalent in the sense that one determines the other, the statistics in Table 1 could all be expressed either as functions of r or $\mathfrak{r}$. T_1, for instance, appears exactly the same in both cases. On the other hand, T_3 is easier to define in terms of $\mathfrak{r}$ and the natural definition of T_5 is in terms of r.

The interpretation of Theorem 1 may be modified to give useful results about rank order tests of independence for bivariate distributions. In the density function $f(x, \theta)$ replace θ by y and now write it in the form $f(x \mid y)$, i.e., the conditional density function of x given y. When X and Y have a joint bivariate distribution assume the conditions of H_1 (and of H_{2k}) are satisfied for $f(x \mid y)$. It is easily verified that when $f(x \mid y)$ satisfies the conditions of H_{2k} then $f(y \mid x)$ satisfies the same conditions. Thus the meaning of "X and Y are jointly of Pólya-type k" is clear.

In the bivariate case define rank order in the following manner: Let $x_1, y_1; \cdots; x_N, y_N$ be N pairs of numbers such that no two of the x's (y's) are equal. Rearrange the order in which the pairs are written to obtain $x_{[1]}, y_{[1]}; \cdots; x_{[N]}, y_{[N]}$ where $y_{[1]} < y_{[2]} < \cdots < y_{[N]}$. The rank order is now given by the vector $r = (r_1, \cdots, r_N)$ where r_i is the number of $x_{[j]} \leqq x_{[i]}$. When the pairs x_i, y_i are replaced by random variables X_i, Y_i the rank order r can be replaced by the corresponding rank order R.

When the pairs X_i, Y_i are mutually independent with a common density function with respect to Lebesgue measure, then the random variable R is defined with probability one. If the null hypothesis (X_i and Y_i are independent) is also true then $\Pr(R = r) = 1/N!$. On the other hand, if X_i and Y_i are jointly of Pólya-type k, then the results of Theorem 2k hold. The proof consists of noting that for given $y_{[1]}, \cdots, y_{[N]}$ the desired inequalities hold between the probabilities of rank orders. Thus the inequalities must hold unconditionally since $m(y) > n(y)$ implies $Em(y) > En(y)$.

The statistics T_1, and T_4 have been frequently proposed as tests of independence and from the results of the preceding paragraph we obtain some evidence that they are admissible when the underlying bivariate distribution is of Pólya-type 2. The other tests in Table 1 could also be used as tests of independence. Those tests which have negative signs in the $r'Lr$ column, however, will be inadmissible.

REFERENCES

[1] D. R. Cox AND A. STUART, "Some quick sign tests for trend in location and dispersion," *Biometrika*, Vol. 42 (1955), pp. 80–95.
[2] F. G. FOSTER AND A. STUART, "Distribution-free tests in time-series based on the breaking of records," *J. Roy. Stat. Soc.* (B), Vol. 16 (1954), pp. 1–22.
[3] S. KARLIN, "Decision theory for Pólya-type distributions—Case of two actions, I," *Third Berkeley Symposium on Probability and Statistics*, Vol. I, pp. 115–129, (1956).
[4] E. L. LEHMANN, "The power of rank tests," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 23–42.
[5] I. R. SAVAGE, "Contributions to the theory of rank orders statistics: the two-sample case," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 590–615.
[6] A. STUART, "The asymptotic relative efficiencies of distribution-free tests of randomness against normal alternatives," *J. Amer. Stat. Assn.*, Vol. 49 (1954), pp. 147–157.
[7] A. STUART, "The efficiencies of tests of randomness against normal regression," *J. Amer. Stat. Assn.*, Vol. 51 (1956), pp. 285–287.