NONPARAMETRIC TESTS AGAINST TREND[1]

By Henry B. Mann

1. INTRODUCTION

The function of a statistical test is usually to decide between two or more courses of action. A test consists in dividing the n-dimensional space into several regions $W_1, \cdots, W_m$ (m may also be infinite). A sample of n measurements $(X_1, \cdots, X_n)$ is then taken and if $(X_1, \cdots, X_n)$ lies in W_i a certain action A_i $(i=1, \cdots, m)$ is taken. The action A_i is the appropriate action if a certain hypothesis H_i is true. In case of two regions we shall say that we test the hypothesis H_1 against the hypothesis H_2.

It is the purpose of the present paper to discuss tests of randomness against trend. In terms of distribution functions the hypothesis H_1 of randomness states that the sample $(X_1, \cdots, X_n)$ is a random sample of n independent variables each with the same continuous distribution function. The hypothesis of a downward trend may be defined in the following way: The sample is still a random sample but X_i has the continuous cumulative distribution function f_i and $f_i(X) < f_{i+k}(X)$ for every i, every X, and every $k > 0$. An upward trend is similarly defined with $f_i(X) > f_{i+k}(X)$.

In testing the hypothesis H_1 of randomness against some class H_2 of alternatives it is customary to fix $P[(X_1, \cdots, X_n) \subset W_1 | H_1]$ where $P(E|H)$ denotes the probability of the event E if H is the true situation. The reason for fixing $P[(X_1, \cdots, X_n) \subset W_1 | H_1]$ is that in this way we can fix the cost of testing, as long as H_1 is the true situation. In quality control, for instance, this means that we fix the cost of controlling a production process that is under statistical control. $1 - P[(X_1, \cdots, X_n) \subset W_1 | H_1]$ is called the size of the critical region.

In proposing a test we usually define for every n a region W_{1n} in the n-dimensional space such that $P[(X_1, \cdots, X_n) \subset W_{1n} | H_1]$ is a fixed constant. Such a test is called consistent with respect to the hypothesis H_2 if, for every alternative B of H_2, $\lim_{n \to \infty} P[(X_1, \cdots, X_n) \subset W_{1n} | B] = 0$.

A test of the hypothesis H_1 against the hypothesis H_2 is called unbiased if for any alternative B of H_2 we have $P[(X_1, \cdots, X_n) \subset W_{1n} | B] \leqq P[(X_1, \cdots, X_n) \subset W_{1n} | H_1)$. Unbiasedness is for all practical purposes a more important requirement than consistency. Suppose, for instance, that a test is biased and an alternative B is true for which $P[(X_1, \cdots, X_n) \subset W_{1n} | B] > P[(X_1, \cdots, X_n) \subset W_{1n} | H_1]$, then the action A_2 is less likely to be taken under the situation B then under

[1] Research under a grant of the Research Foundation of Ohio State University.

TABLE 1*

PROBABILITY OF OBTAINING A PERMUTATION WITH $T \leqq \bar{T}$ IN PERMUTATIONS OF n VARIABLES.

n \ $\bar{T}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
3	167	500	833																			
4	042	167	375	625	833	958																
5	008	042	117	242	408	592	758	883	958	992												
6	001	008	028	068	136	235	360	500	500	640	765	864	932	972	992	999						
7	000	001	005	015	035	068	119	191	281	386	500	500	614	719	809	881	932	965	985	995	999	
8	000	000	001	003	007	016	031	054	089	138	199	274	360	452								
9	000	000	000	000	001	003	006	012	022	038	060	090	130	179	238	306	381	460				
10	000	000	000	000	000	000	001	002	005	008	014	023	036	054	078	108	146	190	242	300	364	431
$P(c)$	000	000	000	000	001	001	002	004	006	010	016	025	037	054	076	105	142	186	237	296	360	429

* Tabular values should be divided by 1000.

$$P(c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-c} e^{-x^2/2} dx, \qquad c = \left(\frac{n(n-1)}{4} - T - \frac{1}{2} \right) \Big/ \sqrt{\frac{2n^3 + 3n^2 - 5n}{72}}, \qquad n = 10.$$

TABLE 2

PROBABILITY OF OBTAINING A PERMUTATION WITH $K \leqq \bar{K}$ IN PERMUTATIONS OF n VARIABLES.

n \ $\bar{K}$	1	2	3	4	5	6	7
3	0.1667	0.5000					
4	0.0417	0.2083	0.5000				
5	0.0083	0.0667	0.2083	0.5000			
6	0.0014	0.0181	0.0792	0.2083	0.5000		
7	0.0002	0.0042	0.0246	0.0792	0.2083	0.5000	
8	0.0000	0.0008	0.0066	0.0284	0.0792	0.2083	0.5000
9	0.0000	0.0002	0.0016	0.0086	0.0284	0.0792	0.2083

the situation H_1 although it should not be taken if H_1 is the true situation. In quality control, for instance, in testing against trend the action A_2 may consist in inspecting machinery to find the causes of a trend. But if a biased test is used then there exist situations when a periodical inspection of machinery would be preferable in deciding the action to be taken. In other words, a biased test is in certain cases not only useless but even worse than no test at all.

In this paper we shall propose two tests against trend and find sufficient conditions for their consistency and unbiasedness. Both tests are based on ranks. The advantages and disadvantages of restricting oneself to tests based on ranks have been discussed in a paper by H. Scheffé.[2] To the advantages of such tests one may add that they may also be used if the quantities considered cannot be measured, as long as it is possible to rank them. Intensity of sensory impressions, pleasure, and pain, are examples of such quantities. In this paper tests against downward trend will be discussed. A test against upward trend can then always be made by testing $-X_1, \cdots, -X_n$ against downward trend.

2. THE T-TEST

Let $X_{i_1}, \cdots, X_{i_n}$ be a permutation of the n distinct numbers $X_1, \cdots, X_n$. Let T count the number of inequalities $X_{i_k} < X_{i_l}$ where $k < l$. One such inequality will be called a reverse arrangement. If $X_1, \cdots, X_n$ all have the same continuous distribution, then the probability of obtaining a sample with T reverse arrangements is proportional to the number of permutations of the variables $1, 2, \cdots, n$ with T reverse arrangements.

The statistic T was first proposed by M. G. Kendall[3] for testing independence in a bivariate distribution. Kendall also derived the recursion formula (1), tabulated the distribution of T for $T \leqq 10$, and proved that the limit distribution of T is normal. Table 1, however, seems more convenient to use for our purpose. The proof of the normality of the limit distribution of T given in the present paper seems simpler than Kendall's and the method employed may be of general interest.

We now propose the following test against a downward trend: We determine $\overline{T}$ so that $P(T \leqq \overline{T} | H_1) = \alpha$, where H_1 is the hypothesis of randomness and α the size of the critical region. If in our sample we obtain a value $T \leqq \overline{T}$ we shall proceed as if the sample came from a

[2] Henry Scheffé, "On a Measure Problem Arising in the Theory of Non-Parametric Tests," *Annals of Mathematical Statistics*, Vol. 14, September, 1943, pp. 227–233.

[3] M. G. Kendall, "A New Measure of Rank Correlation," *Biometrika*, Vol. 30, June, 1938, pp. 81–93.

downward trend. If $T > \overline{T}$ we shall proceed as if H_1 were true. It will be shown that under a mild additional restriction on the sequence $f_1, f_2, \cdots, f_n$ in the alternative H_2 the T-test is a consistent test against trend.

3. THE DISTRIBUTION OF T

Let $P_n(T)$ be the number of permutations of $1, 2, \cdots, n$ with T reverse arrangements. Consider first the permutations of $2, 3, \cdots, n$. We can obtain each permutation of $1, 2, \cdots, n$ exactly once by putting 1 into n different places of all permutations of $2, 3, \cdots, n$. In doing this we increase the number of reverse arrangements by $0, 1, \cdots, n-1$ according to the position into which 1 is placed. Hence

$$(1)\quad P_n(T) = P_{n-1}(T) + P_{n-1}(T-1) + \cdots + P_{n-1}(T-n+1),$$

if $P_n(T) = 0$ for $T < 0$.

Formula (1) permits tabulation of $P_n(T)$ for small values of n. In Table 1 are given the cumulative probabilities of obtaining a permutation with T or fewer reverse arrangements, when every permutation occurs with probability $1/n!$

Since, under the hypothesis of randomness, $P(X_i > X_k) = 1/2$, we have $E(T) = n(n-1)/4$.

To obtain higher moments of T we multiply (1) by $[T - n(n-1)/4]^i = X_n{}^i$. Denoting by $E_n[f(X)]$ the expectation of $f(X_n)$ in permutations of n variables, we obtain

$$(2)\quad \begin{aligned} E_n(X^i) = \frac{1}{n}\Big\{ & E_{n-1}\Big[\Big(X - \frac{n-1}{2}\Big)^i\Big] \\ & + E_{n-1}\Big[\Big(X - \frac{n-3}{2}\Big)^i\Big] + \cdots \\ & + E_{n-1}\Big[\Big(X + \frac{n-1}{2}\Big)^i\Big]\Big\}. \end{aligned}$$

Since the distribution of X is symmetric, $E_n(X^{2i+1}) = 0$, $(i = 0, 1, \cdots)$. From (2) we obtain

$$(3)\quad E_n(X^{2i}) = E_{n-1}(X^{2i}) + \binom{2i}{2} B_n{}^{(2)} E_{n-1}(X^{2i-2}) + \cdots + B_n{}^{(2i)},$$

where

$$(4)\quad B_n{}^{(2j)} = \frac{1}{n} \sum_{k=0}^{k=n-1} \left(\frac{n-1}{2} - k\right)^{2j}.$$

We now put $nB_n{}^{(2j)} = f^{(2j)}(n)$ and $f^{(2j)}(1) = f^{(2j)}(0) = 0$; then $f^{(2j)}(n)$ satisfies for $n = 0, 1, \cdots$, the difference equation

$$f^{(2j)}(n+2) - f^{(2j)}(n) = 2\left(\frac{n+1}{2}\right)^{2j} \tag{5}$$

with the initial condition $f^{(2j)}(1) = f^{(2j)}(0) = 0$. For $j=1$ a solution of (5) is $(n^3-n)/12$. Hence we have $B_n^{(2)} = (n^2-1)/12$. For $j=2$ we obtain $B_n^{(4)} = (3n^4-10n^2+7)/240$. Hence for $i=1$, (3) becomes

$$E_n(X^2) = E_{n-1}(X^2) + \frac{n^2-1}{12}.$$

From this we obtain

$$E_n(X^2) = \sigma_n^2(T) = \frac{2n^3+3n^2-5n}{72}. \tag{6}$$

In a similar manner we obtain

$$E_n(X^4) = \frac{100n^6 + 228n^5 - 455n^4 - 870n^3 + 625n^2 + 372n}{43{,}200}. \tag{7}$$

Formula (6) can also be obtained in a different manner. Let, for $i<j$,

$$y_{ij} = \begin{cases} 1 & \text{if} \quad X_i < X_j, \\ 0 & \text{if} \quad X_i > X_j. \end{cases} \tag{8}$$

Then if the continuous function f is the distribution function of X_i for $i=1, 2, \cdots$, we have

$$E(y_{ij}) = \tfrac{1}{2}, \qquad \sigma^2(y_{ij}) = \tfrac{1}{4},$$
$$E(y_{ij}y_{jk}) = P(X_i < X_j < X_k) = \tfrac{1}{6},$$
$$\sigma(y_{ij}y_{jk}) = -\frac{1}{12}.$$

Similarly we obtain $\sigma(y_{ij}y_{ik}) = \sigma(y_{ij}y_{kj}) = 1/12$, while $\sigma(y_{ij}y_{kl}) = 0$ if i, j, k, l are distinct. For these results (5) can easily be obtained.

We proceed to prove that the limit distribution of T is normal. From (5) and the initial conditions of (5) it follows that $B_n^{(2j)}$ is given by a polynomial in n of degree $2j$. To see this we first determine a polynomial $f^{(2j)}(n)$ of degree $2j+1$ satisfying (5) and the initial condition $f^{(2j)}(0)=0$. It may then be shown by induction that $f^{(2j)}(n) = -f^{(2j)}(-n)$ for all even n. But this is only possible if $f^{(2j)}(n) = -f^{(2j)}(-n)$ for all n. From (5) it follows that $f^{(2j)}(1) = f^{(2j)}(-1)$. Hence we must also have $f^{2j}(1)=0$. Thus there exists a polynomial of degree $2j+1$ satisfying (5) and its initial conditions. We proceed to show by induction that $E_n(X^{2i})$ is given by a polynomial in n of degree $3i$ and first coefficient $[(2i-1)(2i-3)\cdots 3\cdot 1]/36^i$.

From (3) we have

$$(9)\quad \begin{aligned} E_n(X^{2i}) - E_{n-1}(X^{2i}) &= \binom{2i}{2}\frac{(n^2-1)}{12}E_{n-1}(X^{2i-2}) \\ &\quad + \binom{2i}{4}B_n^{(4)}E_{n-1}(X^{2i-4}) + \cdots + B_n^{(2i)}. \end{aligned}$$

From the hypothesis of the induction it follows that the jth term on the right-hand side of (9) is of degree $3i-j$ in n. Hence only the first term is of degree $3i-1$. Since the first difference of $E_n(X^{2i})$ is a polynomial of degree $3i-1$ it follows that $E_n(x^{2i})$ is a polynomial of degree $3i$. Hence we may put $E_n(X^{2i}) = a_0 n^{3i} + \cdots$. Using again the hypothesis of our induction we obtain on comparing coefficients in (9)

$$(10)\quad 3ia_0 = \frac{(2i-3)\cdots 3\cdot 1}{36^{i-1}12}\,\frac{2i(2i-1)}{2}, \quad a_0 = \frac{(2i-1)\cdots 3\cdot 1}{36^i}.$$

From (10) and (6) we have

$$\lim_{n\to\infty}\frac{E_n(X^{2i})}{\sigma_n^{2i}(X)} = (2i-1)\cdots 3\cdot 1.$$

It follows by a well-known theorem that $X/\sigma_n(X)$ is in the limit normally distributed with variance 1. From Table 1 it may be seen that the approach to normality is remarkably rapid.

4. CONDITIONS FOR CONSISTENCY AND UNBIASEDNESS OF THE T-TEST

Let us assume now that some alternative situation (not necessarily a trend) is true. Let, for $i<k$, $P(X_i<X_k) = \frac{1}{2}+\epsilon_{ik}$. Let further y_{ij} be defined by (8) and let $\sum_i\sum_{i<k}\epsilon_{ik} = \lambda_n n(n-1)/2$. Then

$$T = \sum_i\sum_{i<k} y_{ik},$$

$$E(T) = \frac{n(n-1)}{4} + \sum_i\sum_{i<k}\epsilon_{ik} = \frac{n(n-1)}{4}(1+2\lambda_n).$$

Moreover we shall assume that X_i is independent of X_j for $i\neq j$ so that $\sigma(y_{ij}y_{kl}) = 0$ if i, j, k, l are distinct. We proceed to compute $\sigma^2(T)$ under the alternative hypothesis. To simplify the notation the symbol $\sum$ without further specification will denote summation over all values for which the summands have been defined, We have

$$(11)\quad \begin{aligned} \sigma^2(T) = E[(T-E(T))^2] &= \sum \sigma^2(y_{ik}) + 2\sum\sigma(y_{ij}y_{jk}) \\ &\quad + \sum\sigma(y_{kj}\,y_{lj}) + \sum\sigma(y_{ik}y_{il}). \end{aligned}$$

We have

$$\sigma^2(y_{ik}) = \tfrac{1}{4} - \epsilon_{ik}^2, \qquad \sigma(y_{ij}y_{jk}) \leqq 0. \tag{12}$$

The second of these two statements can be proved as follows: Let f_1, f_2, f_3 be three continuous functions, then:

$$\left[\int_{-\infty}^{X_1} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3)\right] \int_{X_1}^{\infty} df_2(X_2)$$
$$\leqq \int_{-\infty}^{X_1} df_2(X_2) \int_{X_1}^{\infty} df_2(X_2) \int_{-\infty}^{X_1} df_3(X_3)$$
$$\leqq \int_{-\infty}^{X_1} df_2(X_2)\left[\int_{X_1}^{\infty} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3)\right].$$

Adding

$$\int_{-\infty}^{X_1} df_2(X_2)\left[\int_{-\infty}^{X_1} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3)\right]$$

to both sides of this inequality we obtain

$$\int_{-\infty}^{X_1} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3) \leqq \int_{-\infty}^{X_1} df_2(X_2) \int_{-\infty}^{+\infty} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3).$$

Integrating both sides with respect to X_1 we obtain

$$\int_{-\infty}^{+\infty} df_1(X_1) \int_{-\infty}^{X_1} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3)$$
$$\leqq \left[\int_{-\infty}^{+\infty} df_1(X_1) \int_{-\infty}^{X_1} df_2(X_2)\right] \left[\int_{-\infty}^{+\infty} df_2(X_2) \int_{-\infty}^{X_2} df_3(X_3)\right]$$

or $P(X_1 > X_2 > X_3) \leqq P(X_1 > X_2)P(X_2 > X_3)$. From this result the inequality in (12) follows easily.

We further have

$$\sigma(y_{ik}y_{il}) = P(y_{ik} = y_{il} = 1) - E(y_{ik})E(y_{il})$$
$$= P(y_{ik} = y_{il} = 1) - (\tfrac{1}{2} + \epsilon_{ik})(\tfrac{1}{2} + \epsilon_{il}),$$

$$P(y_{ik} = y_{il} = 1) = \int_{-\infty}^{+\infty} (1 - f_k)(1 - f_l)df_i \leqq \tfrac{1}{2} + \min(\epsilon_{ik}, \epsilon_{il}).$$

Hence

$$\sigma(y_{ik}y_{il}) \leqq \tfrac{1}{2} + \min(\epsilon_{ik}, \epsilon_{il}) - (\tfrac{1}{2} + \epsilon_{ik})(\tfrac{1}{2} + \epsilon_{il}) \leqq \tfrac{1}{4} - \epsilon_{ik}\epsilon_{il}. \tag{13}$$

Similarly

$$\sigma(y_{kj}y_{lj}) \leqq \tfrac{1}{4} - \epsilon_{kj}\epsilon_{lj}. \tag{14}$$

From (11), (12), (13), and (14) we obtain

$$\sigma^2(T) \leqq \frac{n(n-1)}{4} + \frac{n(n-1)(n-2)}{6} - \sum_{i=1}^{n}\left[\sum_k \epsilon_{ik}^2 + \sum_{k\neq j} \epsilon_{ik}\epsilon_{ij} + \sum_{k\neq j} \epsilon_{ki}\epsilon_{ji}\right].$$

We put $\sum_{k=i+1}^{k=n} \epsilon_{ik} = L_{i\cdot}$, $\sum_{i=1}^{i=k-1} \epsilon_{ik} = L_{\cdot k}$; then

$$\sum_i \left(\sum_k \epsilon_{ik}^2 + \sum_{k\neq j} \epsilon_{ik}\epsilon_{ij} + \sum_{k\neq j} \epsilon_{ki}\epsilon_{ji}\right) = \sum_i L_{i\cdot}^2 + \sum_j L_{\cdot j}^2 - \sum_i\sum_j \epsilon_{ij}^2.$$

Since $\epsilon_{ij}^2 \leqq \frac{1}{4}$ we obtain

$$\sigma^2(T) \leqq \frac{3n(n-1)}{8} + \frac{n(n-1)(n-2)}{6}. \tag{15}$$

If all ϵ_{ij} have the same sign then $L_{i\cdot}^2 \geqq \sum_j \epsilon_{ij}^2$ and $\sum_j L_{\cdot j}^2 \geqq \lambda_n^2 n^2(n-1)/4$. We can then improve (15) to the form

$$\sigma^2(T) \leqq \frac{n(n-1)}{4} + \frac{n(n-1)(n-2)}{6} - \lambda_n^2 \frac{n^2(n-1)}{4}. \tag{15'}$$

If the critical region is given by $T \leqq \overline{T}$ then the power of the T test with respect to the hypothesis H is given by $P(T \leqq \overline{T} | H)$.

If the size of the critical region is fixed then $\overline{T} = n(n-1)/4 - t_n\sqrt{(2n^3+3n^2-5n)/72}$, where $\lim_{n\to\infty} t_n = t$ and $\int_{-\infty}^{-t} e^{-x^2/2}dx$ equals the size of the critical region.

Consider now the case that $\lambda_n < 0$ and $n(n-1)(1+2\lambda_n)/4 < \overline{T}+1$ then by Tchebycheff's theorem we have

$$P(T \leqq \overline{T}|\, H) \geqq 1 - \frac{\sigma^2(T)}{\left[\overline{T} + 1 - \frac{n(n-1)}{4}(1+2\lambda_n)^2\right]}. \tag{16}$$

We may also use

$$P(T \leqq \overline{T}|\, H) \geqq 1 - \frac{\sigma^2(T)}{\left(\lambda_n \frac{n(n-1)}{2} + t_n \sqrt{\frac{2n^3+3n^2-5n}{72}}\right)^2} \tag{16'}$$

for

$$\lambda_n \frac{n(n-1)}{2} \leqq -t_n \sqrt{\frac{2n^3+3n^2-5n}{72}}.$$

For large values of n we may replace t_n by t. From (15), (16′), and the fact that $\lim_{n\to\infty} t_n = t$, it can be seen that the T-test is consistent whenever $\lim_{n\to\infty} \sqrt{n}\lambda_n = -\infty$.

In case the alternative H is a downward trend, we have $f_i(X) < f_j(X)$ if $i<j$ and

$$P(X_i < X_j) = \int (1-f_j)df_i < \int (1-f_i)df_i = \tfrac{1}{2};$$

hence ϵ_{ij} is always negative and (15′) may be used as an upper bound of $\sigma^2(T)$ in (16) or (16′).

Another estimate of $P(T \leqq \overline{T})$, which for small values of n gives better results than (16), can be obtained as follows:

If X is always positive and $E(X)=A$, then

$$P(X > B) = \int_{B+\epsilon}^{\infty} df(X) \leqq \frac{1}{B'}\int_{B+\epsilon}^{\infty} X df(X) \leqq \frac{A}{B'},$$

$$\int_{B+\epsilon}^{\infty} dg = \lim_{\epsilon\to\infty} \int_{B+\epsilon}^{\infty} dg, \qquad \epsilon > 0,$$

where B' is the lower bound of all values B'' for which $\int_{B+\epsilon}^{B''} df(X) > 0$. Thus

$$P(X \leqq B) \geqq 1 - \frac{A}{B'} = \frac{B'-A}{B'}.$$

Hence

$$P(T \leqq \overline{T} \mid H) \geqq \frac{\overline{T} + 1 - \frac{n(n-1)}{4}(1+2\lambda_n)}{\overline{T}+1}. \tag{17}$$

We may also use

$$P(T \leqq \overline{T}) \geqq \frac{-2\lambda_n n(n-1) - 4t_n\sigma_0}{n(n-1) - 4t_n\sigma_0}, \tag{17′}$$

where $\sigma_0 = \sqrt{(2n^3+3n^2-5n)/72}$, and may for large n replace t_n by t. Thus, for instance, for $n=20$, $\lambda_n=0.25$, $t_{20}=1.64$, (17′) yields $P(T \leqq \overline{T}) \geqq 0.326$. A much better result is obtained from (16′) if the distribution of T under the alternative H is approximately normal as is probably the case under a wide class of alternatives.

If the size of the critical region, that is to say, $1-P[(X_1, \cdots, X_n) \subset W_1 | H_1]=\alpha$, then the test is unbiased with respect to H_2 if $P(T \leqq \overline{T} | B) \geqq \alpha$ for every B in H_2. This is, according to (17), the case if

$$\frac{\overline{T}+1-\left(\frac{n(n-1)}{4}+\lambda_n \frac{n(n-1)}{2}\right)}{\overline{T}+1} \geqq \alpha$$

or

$$-\lambda_n \geqq \frac{1}{2}-\frac{(1-\alpha)2(\overline{T}+1)}{n(n-1)}. \tag{18}$$

For instance, if $n=5$, $\alpha=5/120$, then $\overline{T}+1=2$, and we obtain from (18), $-\lambda_n \geqq 0.31$. If n is large enough to use the normal approximation for determining the size of the critical region we obtain with $\sigma_0=\sqrt{(2n^3+3n^2-5n)/72}$

$$-\lambda_n \geqq \frac{\alpha}{2}+\frac{(1-\alpha)t2\sigma_0}{n(n-1)}. \tag{18'}$$

Thus, for example, if $n=10$, $\alpha=0.05$, then $t=1.64$, and we find from (18′) that the T-test is certainly unbiased if $-\lambda_n \geqq 0.218$ [the value obtained from (18) is 0.205]. For $n=20$, $\alpha=0.05$, $t=1.64$, we obtain from (18′), $-\lambda_{20} \geqq 0.154$, which seems satisfactory.

Summary of Section 4: The T-test is consistent with respect to any sequence of random variables $X_1, \cdots, X_n$ for which

1. $P(X_i > X_j)=\frac{1}{2}+\epsilon_{ij}$, for $i<j$;
2. $P(X_i > X_j | X_k > X_l)=P(X_i > X_j)$, if i, j, k, l are distinct;
3. $\lim_{n\to\infty} (\sqrt{n}\sum \epsilon_{ij}/n^2)=-\infty$.

The T-test is unbiased with respect to any set of random variables $X_1, \cdots, X_n$, if $\lambda_n=2\sum \epsilon_{ij}/n(n-1)$ satisfies the inequality (18), which for large n may be replaced by (18′). Lower bounds for the power of the T-test are given by (16), (16′) and (17), (17′) where the primed inequalities are convenient for larger values of n.

5. ALTERNATIVES WITH RESPECT TO WHICH THE T-TEST IS MOST POWERFUL

Let $z_k=i$ if exactly $i-1$ of the X's are larger than X_k. Let $T(i_1, \cdots, i_n)$ be the number of nonreverse arrangements in the permutation $i_1, \cdots, i_n$. $T(i_1, \cdots, i_n)$ is equal to the number of reverse arrangements in the sequence of the X's. Further let us restrict ourselves to tests based on ranks.

Whenever an alternative B is such that $T(i_1, \cdots, i_n) \geqq T(j_1, \cdots, j_n)$ implies $P(z_1=i_1, \cdots, z_n=i_n) \leqq P(z_1=j_1, \cdots, z_n=j_n)$, then the T-test will be most powerful with respect to the alternative B among all tests based on ranks.

We shall as a special case consider a particular alternative B, defined as follows: The probability that, among the set $X_i, X_{i+1}, \cdots, X_n$, X_i will be the first, second, $\cdots$ in magnitude, is, if B is true, given by $a_i, a_ip, \cdots, a_ip^{n-i-1}$ $(p<1)$, $a_i=1/(1+p+\cdots+p^{n-i-1})$ independently of the ranks of the first $i-1$ variables.

Thus, if B is true,

$$P(z_i = k \mid z_1 = j_1, \cdots, z_{i-1} = j_{i-1}) = \begin{cases} 0 & \text{if one } j_\alpha = k, \\ a_ip^{k-l_i-1} & \text{otherwise,} \end{cases}$$

where l_i is the number of j_α's which are $<k$.

Then

$$\begin{aligned} P(z_1 = i_1, z_2 = i_2, \cdots, z_n = i_n) &= a_1p^{i_1-1}a_2p^{i_2-l_2-1} \cdots a_np^{i_n-l_n-1} \\ &= \left(\prod_{i=1}^{i=n} a_i\right)p^{n(n-1)/2}p^{-\Sigma l_i} = \left(\prod_{i=1}^{i=n} a_i\right)p^{T(i_1,\ldots,i_n)}. \end{aligned} \tag{19}$$

Hence $P(z_1=i_1, \cdots, z_n=i_n)<P(z_1=j_1, \cdots, z_n=j_n)$ whenever $T(i_1, \cdots, i_n)>T(j_1, \cdots, j_n)$. Thus the T-test has maximum power with respect to the alternative B. It is, however, not known whether B can result if the X's are independently distributed.

As a side result we obtain from (19) the characteristic function of T. We have

$$\begin{aligned} \prod_{i=1}^{i=n} a_i &= [1(1+p)\cdots(1+p+\cdots+p^{n-1})]^{-1} \\ &= \frac{(p-1)^n}{(p-1)(p^2-1)\cdots(p^n-1)}. \end{aligned}$$

Summing (19) over all permutations, we obtain

$$\frac{1}{n!}\sum P_n(T)p^T = \frac{(p^n-1)(p^{n-1}-1)\cdots(p-1)}{(p-1)^n}.$$

This is an identity in p. Hence $f_T(\theta)$, the characteristic function of T, is given by

$$f_T(\theta) = \sum \frac{1}{n!}P_n(T)e^{iT\theta} = \frac{(e^{in\theta}-1)\cdots(e^{i\theta}-1)}{(e^{i\theta}-1)^n}.$$

6. THE K-TEST

If $P(X_i>X_j)$ increases rapidly with $j-i$, then another test is more powerful than the T-test. This test is carried out as follows:

Determine for the sample $X_0, X_1, \cdots, X_{n-1}$ the smallest value of K for which the following set of inequalities is fulfilled:

$$\begin{aligned} &X_0 > X_k,\ X_0 > X_{k+1}, \cdots, X_0 > X_{n-1}, \\ &X_1 > X_{k+1}, \cdots, X_1 > X_{n-1}, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &X_{n-k-1} > X_{n-1}. \end{aligned} \tag{20}$$

Determine $\bar{K}$ so that $P(K \leqq \bar{K} \mid H_1)$ is equal to the size of the critical region. If $K \leqq \bar{K}$, proceed as if H_2 were true; if $K > \bar{K}$, proceed as if H_1 were true.

Let $X_0, X_1, \cdots, X_{n-1}$ be the sample in the order taken. Consider the $n!$ points $(Y_0, Y_1, \cdots, Y_{n-1})$ where $Y_0 = X_{i_0}, \cdots, Y_{n-1} = X_{i_{n-1}}$ for every permutation $i_0, \cdots, i_{n-1}$ of the numbers $0, 1, \cdots, n-1$. Let $Q_n(\bar{K})$ be the number of points $(Y_0, \cdots, Y_{n-1})$ that satisfy the $\bar{K}$th set of inequalities. We arrange every point in order of decreasing magnitude so that to every point $Y_0, Y_1, \cdots, Y_{n-1}$ we have a sequence of inequalities $Y_{i_0} > \cdots > Y_{i_{n-1}}$. Thus every point is mapped into a permutation $i_0, \cdots, i_{n-1}$. In order that the point with the permutation $i_0, \cdots, i_{n-1}$ fulfill the $\bar{K}$th set of inequalities, it is necessary and sufficient that in the permutation $i_0, i_1, \cdots, i_{n-1}$ no number α be preceded by any of the numbers $\alpha + \bar{K}, \alpha + \bar{K} + 1, \cdots, n-1$. Hence the number of points fulfilling the $\bar{K}$th set of inequalities is equal to the number of permutations in which no number α is preceded by any of the numbers $\alpha + \bar{K}, \alpha + \bar{K} + 1, \cdots, n-1$.

We have $Q_n(1) = 1$. A permutation in which no α is preceded by any number larger than $\alpha + 1$ can have the number 0 only at the first or the second place. Hence such a permutation must either be of the type $0, i_1, \cdots, i_{n-1}$, or of the type $1, 0, i_2, \cdots, i_{n-1}$; and $i_1, \cdots, i_{n-1}$; $i_2, \cdots, i_{n-1}$ respectively must fulfill the $\bar{K}$th set of inequalities with $\bar{K} = 2$. Hence we have the recursion

$$Q_n(2) = Q_{n-1}(2) + Q_{n-2}(2). \tag{21}$$

In using this relation we must put $Q_n(\bar{K}) = 0$ for $n < 0$, $Q_n(n+j) = n!$ for $j \geqq 0$, $n \geqq 0$. To obtain a similar recursion for $\bar{K} = 3$ we observe that a permutation in which α is never preceded by any number larger than $\alpha + 2$ can be only of one of the following types:

$$0, i_1, \cdots, i_{n-1};\ 1, 0, i_2, \cdots, i_{n-1};\ 2, 0, 1, i_3, \cdots, i_{n-1};$$
$$2, 0, 3, 1, i_4, \cdots, i_{n-1};\ 1, 2, 0, i_3, \cdots, i_{n-1};\ 2, 1, 0, i_3, \cdots, i_{n-1}.$$

Hence we have the recursion

$$Q_n(3) = Q_{n-1}(3) + Q_{n-2}(3) + 3Q_{n-3}(3) + Q_{n-4}(3). \tag{22}$$

Let $P_n(\overline{K})$ be the probability of obtaining a permutation satisfying the $\overline{K}$th set of inequalities. For $n \leqq 2\overline{K}$ some of the variables are not involved in the inequalities (20). We therefore have[4]

$$P_n(n - \overline{K}) = P_{2K}(\overline{K}) \quad \text{for} \quad n \geqq 2\overline{K}. \tag{23}$$

We shall show below that $P_8(4)=0.0284$, $P_9(4)=0.0086$, $P_{10}(5)$ $=0.0098$. These values and the relations (21), (22), and (23) permit tabulation of $P_n(\overline{K})$ for $n \leqq 9$. From (23) and Table 2 we further obtain:

$$\begin{aligned} &P_n(n-5) = P_{10}(5) = 0.0098\cdots \quad \text{for} \quad n \geqq 10;\\ &P_n(n-4) = P_8(4) = 0.0284\cdots \quad \text{for} \quad n \geqq 8;\\ &P_n(n-3) = P_6(3) = 0.0792\cdots \quad \text{for} \quad n \geqq 6;\\ &P_n(n-2) = 0.2083\cdots \quad \text{for} \quad n \geqq 4;\\ &P_n(n-1) = 0.5 \quad \text{for} \quad n \geqq 2. \end{aligned} \tag{24}$$

It is clear that (24) contains for $n \geqq 10$ all critical regions possible for the K-test between size 0.0098 and 0.5. Regions smaller than 0.0098 are not likely to occur in practical problems. Hence within a range which is of interest to the practical statistician we shall have all regions available for the K-test if we compute $P_8(4)$, $P_9(4)$, and $P_{10}(5)$. It is a disadvantage of the K-test that we are rather limited in the choice of the size of the critical region.

We shall derive the following two relations: Let $R_n(\overline{K})$ be the subset of the n-dimensional Euclidean space given by (20) and let f be the common cumulative distribution function of the X_i; then for $n \geqq 2\overline{K}$, as we shall prove below,

$$\begin{aligned} P_n(n - \overline{K}) &= P_{2K}(\overline{K}) = \int_{R_{2K}(K)} df(X_0)\cdots df(X_{2K-1})\\ &= \sum \{(j_1+1)[\max(j_1, j_2)+2]\cdots[\max(j_1,\cdots,j_K)+\overline{K}]\}^{-1}, \end{aligned} \tag{25}$$

where $\sum$ denotes summation over all permutations $j_1, \cdots, j_K$ of $1, \cdots, \overline{K}$. Further let $\overline{\max}(i_1, \cdots, i_l) = \min[\max(i_1, \cdots, i_l), \overline{K}-1]$; then

$$\begin{aligned} P_9(4) &= \int_{R_9^{(4)}} df(X_0), \cdots, df(X_8)\\ &= \sum{}' \{[\overline{\max}(j_1)+1][\overline{\max}(j_1,j_2)+2]\\ &\qquad \cdots [\overline{\max}(j_1,\cdots,j_5)+5]\}^{-1}, \end{aligned} \tag{26}$$

[4] For typographical reasons it was not possible to use $\overline{K}$ in subscripts and exponents. Every K in subscripts and exponents in this section should be read as $\overline{K}$.

where $\sum'$ denotes summation over all permutations $j_1, \cdots, j_5$ of $1, \cdots, 5$ for which 1 precedes 5.

Since the integral in (25) is independent of f, we may assume that f is a rectangular distribution between 0 and 1. We consider the integral in (25). For $X_0, X_1, \cdots, X_{K-1}$ fixed, X_K varies from 0 to X_0; X_{K+1} from 0 to $\min(X_0, X_1)$; X_{K+2} from 0 to $\min(X_0, X_1, X_2)$; and so forth; hence we obtain

$$P_{2K}(\overline{K}) = \int_0^1 \cdots \int_0^1 X_0 \min(X_0, X_1) \cdots \min(X_0, X_1, \cdots, X_{K-1}) dX_0 \cdots dX_{K-1}.$$

We split this integral into the $\overline{K}!$ parts $X_{i_1} < X_{i_2} < \cdots < X_{i_K}$; then for any permutation $i_1, \cdots, i_K$ of $0, \cdots, \overline{K}-1$ we have to compute

$$\int_0^1 dX_{i_K} \int_0^{X_{i_K}} dX_{i_{K-1}} \cdots \int_0^{X_{i_2}} dX_{i_1} X_0 \min(X_0, X_1) \cdots \min(X_0, \cdots, X_{K-1}). \tag{27}$$

Consider the exponent of X_{i_α} in (27). In the factors under the integral sign X_{i_α} occurs for the first time in $\min(X_0, X_1, \cdots, X_{i_\alpha})$. From then on it occurs in every factor. If $i_\beta < i_\alpha$ for $\beta < \alpha$, then none of the factors in (27) will be equal to X_{i_α}. If $\min(i_1, \cdots, i_{\alpha-1}) = i_\beta > i_\alpha$ $(\alpha > 1)$, then the factors $\min(X_0, \cdots, X_{i_\alpha}), \min(X_0, \cdots, X_{i_\alpha+1}), \cdots, \min(X_0, \cdots, X_{i_\beta-1})$ will be equal to X_{i_α}. In both cases the exponent of X_{i_α} may be written as $\min(i_1, \cdots, i_{\alpha-1}) - \min(i_1, \cdots, i_\alpha)$. For $\alpha = 1$, we shall have $\min(X_0, \cdots, X_{i_1}) \cdots \min(X_0, \cdots, X_{K-1})$ equal to X_{i_1}. Hence (27) becomes

$$\int_0^1 dX_{i_K} X_{i_K}^{\min(i_1, \cdots, i_{K-1}) - \min(i_1, \cdots, i_K)} \cdot \int_0^{X_{i_K}} \cdots \int_0^{X_{i_3}} dX_{i_2} X_{i_2}^{i_1 - \min(i_1, i_2)} \int_0^{X_{i_2}} dX_{i_1} X_{i_1}^{K - i_1}. \tag{28}$$

Integrating out the last integral we obtain under the next to the last integral sign

$$\frac{1}{\overline{K} - i_1 + 1} X_{i_2}^{K - i_1 + 1} X_{i_2}^{i_1 - \min(i_1, i_2)} = (\overline{K} - i_1 + 1)^{-1} X_{i_2}^{K - \min(i_1, i_2) + 1}.$$

When this process is continued (28) finally becomes

$$(\overline{K} - i_1 + 1)^{-1} [\overline{K} - \min(i_1, i_2) + 2]^{-1} \cdots [\overline{K} - \min(i_1, \cdots, i_K) + \overline{K}]^{-1}. \tag{29}$$

Putting $\bar{K}-i_\alpha=j_\alpha$ and summing over all permutations $i_1, \cdots, i_K$, we obtain (25).

In the integral in (26) we have X_5 varying from 0 to min (X_0, X_1); X_6 from 0 to min (X_0, X_1, X_2); and so forth, hence

$$P_9(4) = \int_0^1 \cdots \int_0^1 \int_0^{X_0} \min (X_0, X_1) \cdots \min (X_0, X_1, \cdots, X_4) dX_4 \cdots dX_0. \tag{30}$$

Now for every subset $X_{i_1} < \cdots < X_{i_5}$ where $i_1, \cdots, i_5$ is a permutation of the numbers 0, 1, 2, 3, 4, we obtain

$$\int_0^1 \int_0^{X_{i_4}} \cdots \int_0^{X_{i_2}} \min (X_0, X_1) \cdots \min (X_0, \cdots, X_4) dX_{i_1} \cdots dX_{i_5}$$

$$= \int_0^1 \int_0^{X_{i_4}} \cdots \int_0^{X_{i_2}} X_{i_1}{}^{5-\max (i_1,1)} X_{i_2}{}^{\max (i_1,1)-\max [\min (i_1,i_2),1]} \cdots X_{i_5}{}^{\max [\min (i_1,i_2,i_3,i_4),1]-\max [\min (i_1,i_2,i_3,i_4,i_5),1]} dX_{i_1} \cdots dX_{i_5},$$

from which (26) follows by an obvious extension of the argument used in the proof of (25).

By the use of (25) and (26) the values $P_8(4)$, $P_{10}(5)$, and $P_9(4)$ have been computed. It may be remarked that the labor involved was not at all excessive since in carrying out the computations a great many short cuts offer themselves freely.

It is easy to construct trends for which the K-test has maximum power at some fixed level of significance. Let the critical value of K be $\bar{K}$ and consider alternatives where $P(X_i > X_{\bar{K}+i+j}) = 1$ for $j \geqq 0$. It is easy to find trends for which this is true. The K-test has then clearly the power 1 with respect to such trends. The power of the K-test is also high with respect to trends for which $P(X_1 > X_{i+\bar{K}+j})$ is close to 1 for $j \geqq 0$. Such alternatives do quite frequently occur in practical work. The fact that the K-test is most powerful with respect to a fairly wide class of alternatives seems worth noting. It is very often said that in using a test based on ranks, one is "throwing away information." In using the K-test an even larger "amount of information" is "thrown away," nevertheless it is a most powerful test with respect to a substantial class of alternatives.

The Ohio State University