ON A TEST OF WHETHER ONE OF TWO RANDOM VARIABLES IS STOCHASTICALLY LARGER THAN THE OTHER

By H. B. Mann and D. R. Whitney

Ohio State University

1. Summary. Let x and y be two random variables with continuous cumulative distribution functions f and g. A statistic U depending on the relative ranks of the x's and y's is proposed for testing the hypothesis $f = g$. Wilcoxon proposed an equivalent test in the *Biometrics Bulletin*, December, 1945, but gave only a few points of the distribution of his statistic.

Under the hypothesis $f = g$ the probability of obtaining a given U in a sample of n x's and m y's is the solution of a certain recurrence relation involving n and m. Using this recurrence relation tables have been computed giving the probability of U for samples up to $n = m = 8$. At this point the distribution is almost normal.

From the recurrence relation explicit expressions for the mean, variance, and fourth moment are obtained. The $2r$th moment is shown to have a certain form which enabled us to prove that the limit distribution is normal if m, n go to infinity in any arbitrary manner.

The test is shown to be consistent with respect to the class of alternatives $f(x) > g(x)$ for every x.

2. Introduction. Let x and y be two random variables having continuous cumulative distribution functions f and g respectively. The variable x will be called stochastically smaller than y if $f(a) > g(a)$ for every a. We wish to test the hypothesis $f = g$ against the alternative that x is stochastically smaller than y. Such alternatives are of great importance in testing, for instance, the effect of treatments on some measurement. One may think of x as the values of certain measurements in the control group and of y as the values of the same measurement in a group which received treatment. In a particular instance the protective effect against infection by certain bacteria was investigated. Two groups of rats were used in the experiment. The first group receiving no treatment, the second group receiving the drug. Both groups were then infected with supposedly equally diluted cultures of the bacteria under investigation. Most of the rats in both groups died, but the time of survival was measured and it was desired to test whether the drug had the effect of prolonging the life of the rats. It was desired to make inferences from the effect on rats to the effect the drug would have on humans. Thus, the only relevant alternative to the hypothesis that survival times are not influenced by the drug is that the survival time of those rats which received treatment is stochastically larger than that of the control group.

3. The U test. Let the quantities $x_1, \cdots, x_n, y_1, \cdots, y_m$ be arranged in order. This arrangement is unique with probability 1 if $P(x_i = y_j) = 0$ and this follows from our assumption of continuity. Let U count the number of times a y precedes an x. If $P(U \leq \overline{U}) = \alpha$ under the null hypothesis, the test will be considered significant on the significance level α if $U \leq \overline{U}$ and the hypothesis of identical distributions of x and y will be rejected.

This test was first proposed by Wilcoxon [1]. His statistic T is the sum of the ranks of the y's in the ordered sequence of x's and y's. In general

$$U = mn + \frac{m(m+1)}{2} - T$$

and this gives a simple way of computing U. Wilcoxon, however, treated only the case $m = n$ and in this case he tabulated only 3 points of the distribution of T. Since the test seems of great utility it seemed worthwhile to compute the variance, the moments and the limit distribution of U and to investigate the class of alternatives with respect to which the test is consistent.

Although this paper is written in terms of U and the probabilities of U are tabulated the results can be easily interpreted in terms of T if so desired.

4. The distribution of U. Consider now ordered sequences of n x's and m y's. Since it is only the relation between x and y that matters we replace each x by a 0 and each y by a 1. Let U count the number of times a 1 precedes a 0. Let $\bar{p}_{nm}(U)$ be the number of sequences of n 0's and m 1's in each of which a 1 precedes a 0 U times. By examining a sequence with the last term omitted we arrive at the recurrence relation:

$$\bar{p}_{nm}(U) = \bar{p}_{n-1m}(U - m) + \bar{p}_{nm-1}(U),$$

where $\bar{p}_{ij}(U) = 0$ if $U < 0$ and $\bar{p}_{i0}(U)$, $\bar{p}_{0i}(U)$ are zero or one according as $U \neq 0$ or $U = 0$.

Under the null hypothesis each of the $(m + n)!/m!n!$ sequences of n 0's and m 1's is equally likely. Consequently if $p_{nm}(U)$ represents the probability of a sequence in which a 1 precedes a 0 U times then

$$p_{nm}(U) = \frac{n}{n+m} p_{n-1m}(U - M) + \frac{m}{n+m} p_{nm-1}(U). \tag{1}$$

Using the recurrence relation (1) the probabilities $p_{nm}(U)$ have been tabulated for $m \leq n \leq 8$ (see Table I). For $m = n = 8$ the distribution of $U - \frac{1}{2}(nm + 1)$ differs only a negligible amount from the normal distribution. We shall, in the following, derive the mean, the variance, and the fourth moment of U, and prove that the limit distribution of U is normal if n and m both approach infinity in any arbitrary manner.

It is obvious that $p_{nm}(U) = p_{mn}(U)$.

Since the probability of the ith 1 preceding the jth 0 is $\frac{1}{2}$, we have

$$E_{nm}(U) = nm/2. \tag{2}$$

TABLE I

Probability of Obtaining a U not Larger than that Tabulated in Comparing Samples of n and m

$n = 3$

U \ m	1	2	2
0	.250	.100	.050
1	.500	.200	.100
2	.750	.400	.200
3		.600	.350
4			.500
5			.650

$n = 4$

U \ m	1	2	3	4
0	.200	.067	.028	.014
1	.400	.133	.057	.029
2	.600	.267	.114	.057
3		.400	.200	.100
4		.600	.314	.171
5			.429	.243
6			.571	.343
7				.443
8				.557

$n = 5$

U \ m	1	2	3	4	5
0	.167	.047	.018	.008	.004
1	.333	.095	.036	.016	.008
2	.500	.190	.071	.032	.016
3	.667	.286	.125	.056	.028
4		.429	.196	.095	.048
5		.571	.286	.143	.075
6			.393	.206	.111
7			.500	.278	.155
8			.607	.365	.210
9				.452	.274
10				.548	.345
11					.421
12					.500
13					.579

$n = 6$

U \ m	1	2	3	4	5	6
0	.143	.036	.012	.005	.002	.001
1	.286	.071	.024	.010	.004	.002
2	.428	.143	.048	.019	.009	.004
3	.571	.214	.083	.033	.015	.008
4		.321	.131	.057	.026	.013
5		.429	.190	.086	.041	.021
6		.571	.274	.129	.063	.032
7			.357	.176	.089	.047
8			.452	.238	.123	.066
9			.548	.305	.165	.090
10				.381	.214	.120
11				.457	.268	.155
12				.545	.331	.197
13					.396	.242
14					.465	.294
15					.535	.350
16						.409
17						.469
18						.531

TABLE I (*Continued*)

$n = 7$

U \ m	1	2	3	4	5	6	7
0	.125	.028	.008	.003	.001	.001	.000
1	.250	.056	.017	.006	.003	.001	.001
2	.375	.111	.033	.012	.005	.002	.001
3	.500	.167	.058	.021	.009	.004	.002
4	.625	.250	.092	.036	.015	.007	.003
5		.333	.133	.055	.024	.011	.006
6		.444	.192	.082	.037	.017	.009
7		.556	.258	.115	.053	.026	.013
8			.333	.158	.074	.037	.019
9			.417	.206	.101	.051	.027
10			.500	.264	.134	.069	.036
11			.583	.324	.172	.090	.049
12				.394	.216	.117	.064
13				.464	.265	.147	.082
14				.538	.319	.183	.104
15					.378	.223	.130
16					.438	.267	.159
17					.500	.314	.191
18					.562	.365	.228
19						.418	.267
20						.473	.310
21						.527	.355
22							.402
23							.451
24							.500
25							.549

TABLE I (*Continued*)

$n = 8$

U \ m	1	2	3	4	5	6	7	8	t	normal
0	.111	.022	.006	.002	.001	.000	.000	.000	3.308	.001
1	.222	.044	.012	.004	.002	.001	.000	.000	3.203	.001
2	.333	.089	.024	.008	.003	.001	.001	.000	3.098	.001
3	.444	.133	.042	.014	.005	.002	.001	.001	2.993	.001
4	.556	.200	.067	.024	.009	.004	.002	.001	2.888	.002
5		.267	.097	.036	.015	.006	.003	.001	2.783	.003
6		.356	.139	.055	.023	.010	.005	.002	2.678	.004
7		.444	.188	.077	.033	.015	.007	.003	2.573	.005
8		.556	.248	.107	.047	.021	.010	.005	2.468	.007
9			.315	.141	.064	.030	.014	.007	2.363	.009
10			.387	.184	.085	.041	.020	.010	2.258	.012
11			.461	.230	.111	.054	.027	.014	2.153	.016
12			.539	.285	.142	.071	.036	.019	2.048	.020
13				.341	.177	.091	.047	.025	1.943	.026
14				.404	.217	.114	.060	.032	1.838	.033
15				.467	.262	.141	.076	.041	1.733	.041
16				.533	.311	.172	.095	.052	1.628	.052
17					.362	.207	.116	.065	1.523	.064
18					.416	.245	.140	.080	1.418	.078
19					.472	.286	.168	.097	1.313	.094
20					.528	.331	.198	.117	1.208	.113
21						.377	.232	.139	1.102	.135
22						.426	.268	.164	.998	.159
23						.475	.306	.191	.893	.185
24						.525	.347	.221	.788	.215
25							.389	.253	.683	.247
26							.433	.287	.578	.282
27							.478	.323	.473	.318
28							.522	.360	.368	.356
29								.399	.263	.396
30								.439	.158	.437
31								.480	.052	.481
32								.520		

We now seek an expression for $E_{nm}(u^2)$ where $u = U - nm/2$. After multiplying (1) by $(U - nm/2)^2$, using

$$E_{nm}(u^2) = \sum_U (U - nm/2)^2 p_{nm}(U)$$

and expanding:

$$E_{nm}(u^2) = \frac{n}{n+m} E_{n-1m}(u^2) + \frac{m}{n+m} E_{nm-1}(u^2) + nm/4, \tag{3}$$

where $E_{nm}(u)$ denotes the expectation of $(U - nm/2)$ in sequences with n 0's and m 1's. The initial conditions of (3) are seen by direct calculation to be

$$E_{n0}(u^2) = E_{0m}(u^2) = 0. \tag{4}$$

By substitution $E_{nm}(u^2) = nm(n + m + 1)/12$ is a solution of the recurrence relation (3) and its initial conditions (4). Hence, it follows by mathematical induction that

$$E_{nm}(u^2) = nm(n + m + 1)/12. \tag{5}$$

The fourth moment is similarly a solution of the recurrence relation

$$E_{nm}(u^4) = \frac{n}{n+m} E_{n-1m}(u^4) + \frac{m}{n+m} E_{nm-1}(u^4) + \frac{nm}{16}(2n^2m + 2nm^2 - n^2 - m^2 - nm) \tag{6}$$

which is obtained from (1) by multiplication by $(U - nm/2)^4$ and expansion. The initial conditions of (6) are found by direct calculation to be

$$E_{n0}(u^4) = E_{0m}(u^4) = 0. \tag{7}$$

It may be verified that

$$E_{nm}(u^4) = \frac{nm(n+m+1)}{240}(5n^2m + 5nm^2 - 2n^2 - 2m^2 + 3nm - 2n - 2m) \tag{8}$$

satisfies the recurrence relation (6) and its initial conditions (7) and hence (8) follows by mathematical induction.

To investigate the limit distribution of u as n, m become infinite we investigate the rth moment. Following the same procedure as in the case of the second and fourth moments and using the symmetry of the distribution to find the odd moments zero we get the following recurrence relation.

$$E_{nm}(u^{2r}) = \frac{1}{n+m}\sum_{\alpha=0}^{r}\binom{2r}{2\alpha}\frac{1}{4^\alpha}\{nm^{2\alpha}E_{n-1m}(u^{2r-2\alpha}) + mn^{2\alpha}E_{nm-1}(u^{2r-2\alpha})\} \tag{9}$$

For $r = 1, 2$ it is known that $E_{nm}(u^{2r})$ is a polynomial in n and m of degree $3r$ and that it is divisible by $nm(n + m + 1)$. Assuming that $E_{nm}(u^{2\alpha})$, $\alpha < r$ is a polynomial in n and m of degree 3α divisible by $nm(n + m + 1)$ we will show that it is possible to find a polynomial of degree $3r$ in n and m divisible by $nm(n + m + 1)$ which satisfies the recurrence relation (9) for $E_{nm}(u^{2r})$ and also its initial conditions, namely, $E_{n0}(u^{2r}) = E_{0m}(u^{2r}) = 0$.

The last condition is trivially satisfied if $E_{nm}(u^{2r})$ is divisible by $nm(n + m + 1)$. Our method here is to actually substitute a polynomial with undetermined coefficients into (9) and show that the coefficients can be obtained uniquely. Rearranging (9) we obtain

$$\begin{aligned}(10)\qquad & E_{nn}(u^{2r}) - \frac{n}{n+m} E_{n-1m}(u^{2r}) - \frac{m}{n+m} E_{nm-1}(u^{2r}) \\ & = \frac{1}{n+m}\sum_{\alpha=1}^{r}\binom{2r}{2\alpha}\frac{1}{4^{\alpha}}\{nm^{2\alpha} E_{n-1m}(u^{2r-2\alpha}) + mn^{2\alpha} E_{nm-1}(u^{2r-2\alpha})\}\end{aligned}$$

Since for $\lambda < r$ we can write $E_{nm}(u^{2\lambda}) = nm(n + m + 1)P_{nm}^{3\lambda-3}$ where $P_{nm}^{3\lambda-3}$ is a polynomial in n, m of degree $3\lambda - 3$ the above equation reduces to

$$(11)\qquad E_{nm}(u^{2r}) - \frac{n}{n+m} E_{n-1m}(u^{2r}) - \frac{m}{n+m} E_{nm-1}(u^{2r}) = nmQ_{nm}^{3r-3}$$

where Q_{nm}^{3r-3} is a polynomial in n, m of degree $3r - 3$.

Now let

$$E_{nm}(u^{2r}) = nm(n + m + 1) \sum_{\substack{i,j=0 \\ i+j\leq 3r-3}}^{3r-3} a_{ij} n^{i} m^{j}$$

where $a_{ij} = a_{ji}$ are to be determined. Substitution in (11) yields:

$$\sum_{i,j} a_{ij}[(n + m + 1)n^{i}m^{j} - (n - 1)(n - 1)^{i} m^{j} - (m - 1)n^{i}(m - 1)^{j}] = Q_{nm}^{3r-3}$$

and rearrangement yields:

$$(12)\qquad \sum_{\substack{j,j=0 \\ i+j\leq 3r-3}}^{3r-3} a_{ij}\left[n^{i}m^{j} + \sum_{\alpha=0}^{i}\binom{i+1}{\alpha}(-1)^{i-\alpha}(n^{j}m^{\alpha} + n^{\alpha}m^{j})\right] = Q_{nm}^{3r-3}.$$

Consider first the terms of degree $3r - 3$. In this case $i + j = 3r - 3$ and $\alpha = i$ will give

$$\sum_{i=0}^{3r-3} a_{i3r-3-i}[n^{i}m^{3r-3-i} + (i + 1)(n^{3r-3-i}m^{i} + n^{i}m^{3r-3-i})]$$

or

$$(13)\qquad 3r\sum_{i=0}^{3r-3} a_{i3r-3-i}\, n^{i} m^{3r-3-i}.$$

Equating the coefficients of these terms of degree $3r - 3$ to the corresponding ones in Q_{nm}^{3r-3} it is possible to calculate the value of $a_{i3r-3-i}$, $(i = 0, \cdots, 3r - 3)$.

We assume now that the a_{ij} are known for $i + j \geq 3r - 3 - (k - 1)$ and we will find the value of a_{ij} where $i + j = 3r - 3 - k$. Consider then the terms in (12) of degree $3r - 3 - k$. These terms will occur when

$$i + j = 3r - 3, \alpha = i - k; i + j = 3r - 4, \alpha = i - k + 1; \cdots ; \quad i + j = 3r - 3 - k, \alpha = i.$$

All, but the last, contain coefficients which have already been evaluated. The last one reduces to

$$(3r - 3) \sum_{i=0}^{3r-3-k} a_{i3r-3-k-i}\, n^i m^{3r-3-k-i}.$$

Thus by equating coefficients $a_{i3r-3-k-i}$ for $i = 0, 1, \cdots 3r - 3 - k$ can be evaluated in terms of the coefficients a_{ij} already known and those in Q_{nm}^{3r-3}. This concludes the proof that $E_{nm}(u^{2r})$ is a polynomial in n, m of degree $3r$ and is divisible by $nm(n + m + 1)$.

We now investigate the coefficients of the terms of degree $3r$. For $\lambda = 1, 2$

$$E_{nm}(u^{2\lambda}) = \frac{(2\lambda - 1) \cdots 5\cdot3\cdot1}{12^\lambda} (nm)^\lambda (n + m + 1)^\lambda + \text{terms of degree} < 3\lambda).$$

We assume this to hold for $\lambda < r$ and we will show that it holds for $\lambda = r$. Substitution reduces the right side of (10) to

$$\frac{1}{n+m}\binom{2r}{2}\frac{1}{4}\left\{nm^2\left[\frac{(2r-3)\cdots5\cdot3\cdot1}{12^{r-1}}(n-1)^{r-1}m^{r-1}(n+m)^{r-1}\right]\right.$$
$$\left.+ mn^2\left[\frac{(2r-3)\cdots5\cdot3\cdot1}{12^{r-1}}n^{r-1}(m-1)^{r-1}(n+m)^{r-1}\right] + (\text{terms of degree} < 3r)\right\}$$

or

$$\frac{r(2r-1)}{4}\left\{(n+m)^{r-2}\left[\frac{(2r-3)\cdots5\cdot3\cdot1}{12^{r-1}}\right]\right.$$
$$\left.\cdot [n(n-1)^{r-1}m^{r+1} + m(m-1)^{r-1}n^{r+1}] + (\text{terms of degree} < 3r - 1)\right\}$$

which reduces to

$$\frac{3r(2r-1)\cdots5\cdot3\cdot1}{12^r}(nm)^r(n+m)^{r-1} + (\text{terms of degree} < 3r - 1).$$

Comparison of coefficients with (13) multiplied by nm gives

$$nm\sum_{i=0}^{3r-3} a_{i3r-3-i}\, n^i m^{3r-3-i} = \frac{(2r-1)\cdots5\cdot3\cdot1}{12^r}(nm)^r(n+m)^{r-1}$$

or

$$E_{nm}(u^{2r}) = \frac{(2r-1)\cdots5\cdot3\cdot1}{12^r}(nm)^r(n+m+1)^r + (\text{terms of degree} < 3r). \tag{14}$$

We now wish to show that $E_{nm}(u^{2r})$ is at most of degree $2r$ in n or m. For $r = 1, 2$ this has already been established. Assuming that it is true for lower moments the right side of (10), which reduces to nmQ_{nm}^{3r-3} is at most of degree $2r - 1$ in n. We again compare coefficients in (12). First, for terms of degree $3r - 3$ we have already seen that n has degree at most $2r - 2$. For terms of degree $3r - 4$ we use $i + j = 3r - 3$, $\alpha = i - 1$ and $i + j = 3r - 4$, $\alpha = i$. The first case gives rise to no terms in n of degree greater than $2r - 2$ so when we solve for the coefficients $a_{i3r-4-i}$ the coefficients of terms in n of degree greater than $2r - 2$ must be zero. The process repeats and we find no terms in n or m of degree greater than $2r - 2$ in the left side of (12). This gives $E_{nm}(u^{2r})$ at most the degree $2r$ in n or m.

Now consider the ratio

$$I \equiv \frac{E_{nm}(u^{2r})}{[E_{nm}(u^2)]^r}$$

$$= \frac{\frac{(2r-1)\cdots 5\cdot 3\cdot 1}{12^r}(nm)^r(n+m+1)^r}{[nm(n+m+1)/12]^r} + \frac{(\text{terms of degree} < 3r;\text{ in } n \text{ or } m, \leq 2r)}{[nm(n+m+1)/12]^r}$$

$$= (2r-1)\cdots 5\cdot 3\cdot 1 + \frac{(\text{terms of degree} < 3r;\text{ in } n \text{ or } m, \leq 2r)}{(nm)^r(n+m+1)^r}.$$

Hence

$$\lim_{n,m\to\infty} I = (2r-1)\cdots 5\cdot 3\cdot 1 \tag{15}$$

and by a well known theorem it follows from (15) that the limit distribution is normal.

5. Consistency of the U test. If f and g are the cumulative distribution functions of the x's and y's then our null hypothesis is $f = g$. The alternatives admitted are $f(a) > g(a)$ for every a. Let E_A denote the expectation under the alternative.

Defining

$$x_{ij} = \begin{matrix} 0 \text{ if } x_i < y_j \\ 1 \text{ if } x_i > y_j \end{matrix}$$

we have

$$E_A(x_{ij}) = P(x_i > y_j) = \int_{-\infty}^{\infty} g\, df < \tfrac{1}{2}$$

$$E_A(x_{ij}x_{ik}) = P(x_i > y_j;\, x_i > y_k) = \int_{-\infty}^{\infty} g^2\, df < \tfrac{1}{3}$$

$$E_A(x_{ik}x_{jk}) = P(x_i > y_k,\, x_j > y_k) = \int_{-\infty}^{\infty} (1-f)^2\, dg < \tfrac{1}{3}.$$

We can now write

$$E_A(x_{ij}) = \tfrac{1}{2} - \lambda, \qquad E_A(x_{ij}x_{ik}) = \tfrac{1}{3} - \epsilon_1, \qquad E_A(x_{ik}x_{jk}) = \tfrac{1}{3} - \epsilon_2$$

where λ, ϵ_1, ϵ_2 are positive numbers.
We have then

$$\sigma_A^2(x_{ij}) = \tfrac{1}{4} - \lambda^2 \qquad \sigma_A(x_{ij}x_{ik}) = \tfrac{1}{12} - \epsilon_1 + \lambda - \lambda^2$$

$$\sigma_A(x_{ij}x_{kl}) = 0 \text{ for } i \neq k, j \neq l \quad \sigma_A(x_{ik}x_{jk}) = \tfrac{1}{12} - \epsilon_2 + \lambda - \lambda^2$$

Now

$$E_A(U) = \sum_{i,j} E_A(x_{ij}) = nm/2 - \lambda nm \tag{16}$$

and

$$\sigma_A^2(U) = \Sigma\, \sigma^2{}_A(x_{ij}) + \Sigma\, \sigma_A(x_{ij}x_{ik}) + \Sigma\, \sigma_A(x_{ik}x_{jk}) + \Sigma\, \sigma_A(x_{ij}x_{kl}) \tag{17}$$

or

$$\sigma_A^2(U) = nm(n + m + 1)/12 + nm[-\lambda^2(n + m - 1) + (\lambda - \epsilon_1)(m - 1) + (\lambda - \epsilon_2)(n - 1)].$$

Let the critical region under the null hypothesis consist of those U's satisfying $nm/2 - U \geq t_n\sigma$ where $\lim_{n\to\infty} t_n = t$. Then

$$P(nm/2 - U \geq t_n\sigma \mid A) = P(E_A(U) - U \geq k\cdot\sigma_A) \quad \text{where} \quad k = \frac{t_n\sigma - \lambda nm}{\sigma_A}$$

and by Tchebycheff's inequality, since for large values of n,m $k < 0$

$$P(nm/2 - U \geq t_n\sigma \mid A) \geq 1 - \frac{\sigma_A^2}{(t_n\sigma - \lambda nm)^2},$$

which by (5) and (17) gives

$$P(nm/2 - U \geq t_n\sigma \mid A) \geq 1 - \frac{\dfrac{nm(n+m+1)}{12} + nm[-\lambda^2(n+m-1) + (\lambda - \epsilon_1)(m-1) + (\lambda - \epsilon_2)(n-1)}{(t_n\sqrt{nm(n+m+1)/12} - \lambda nm)^2}$$

$$\geq 1 - \frac{1 + \dfrac{12}{n+m+1}[-\lambda^2(n+m-1) + (\lambda - \epsilon_1)(m-1) + (\lambda - \epsilon_2)(n-1)]}{\left(t_n - \lambda\sqrt{\dfrac{12nm}{n+m+1}}\right)^2}.$$

We obtain then that

$$\operatorname*{Lim}_{n,m\to\infty} P(nm/2 - U \geq t_n\sigma \mid A) = 1$$

which is the requirement for consistency.

6. Comparison with other tests. Another test which might seem appropriate for the comparison of a control group with a group receiving treatment is the test introduced by Wald and Wolfowitz [2]. The test by Wald and Wolfowitz is consistent with respect to every alternative g. However in the case considered we are only interested in the alternative hypothesis that measurements in the group receiving treatment are stochastically larger than in the control group. Intuitively, it seems that the test proposed here is more efficient for detecting the particular alternative considered than the test proposed by Wald and Wolfowitz. This intuitive feeling was borne out by the results of the test in the particular experiment described in the introduction. All in all, 62 experiments were conducted using various bacteria in different solutions and various amounts of the protective drug. The U Test gave 14 significant results on the 5% level and 4 on the 1% level. The test of Wald and Wolfowitz gave 7 significant results on the 5% level and 2 on the 1% level. A final decision between the two tests can, of course, only be arrived at on the basis of their power functions, which present formidable difficulties.

In comparing the two statistics it was noted that a slight dislocation of a value may cause a significant change in the number of runs easier than it can cause a significant change in the statistic proposed here. For instance, in the sequence $x_1x_2x_3x_4x_5x_6y_1y_2y_3y_4y_5y_6$ both statistics would give a probability less than .05. If however, the sequence is slightly altered to $x_1x_2x_3x_4x_5y_1x_6y_2y_3y_4y_5y_6$, P (number of runs ≤ 4) $> .05$ while $P(U \leq 1) = .002$.

After completion of the present paper it came to the authors attention that the U test had already been proposed by K. K. Mathen [3]. However Mathen's distribution of U is incorrect and its derivation erroneous, since it assumes independence of the random variables x_{ij} as defined in section 5 of the present paper, while obviously x_{ij} and x_{ik} are not independent.

REFERENCES

[1] Frank Wilcoxon, "Individual comparisons by ranking methods", *Biometrics Bull.*, Vol. 1 (1945), pp. 80–83.

[2] A. Wald and J. Wolfowitz, "On a test whether two samples are from the same population", *Annals of Math. Stat.*, Vol. 11 (1940), pp. 147–162.

[3] K. K. Mathen, *Sankhyā*, 1946, p. 329.