September 1966

ON A GENERALIZATION OF WILCOXON'S RANK SUM TEST FOR CENSORED DATA

Milton Sobel

Technical Report No. 69 (Revised)

University of Minnesota
Minneapolis, Minnesota

On a Generalization of Wilcoxon's

Rank Sum Test for Censored Data*

by Milton Sobel

University of Minnesota

1. Summary

A statistic is proposed for a two sample unpaired nonparametric test similar to Wilcoxon's test but more suitable for censored data; in life-testing applications the data are times of failure. There are m + n units put on test at the outset, m from one population and n from the other. In the problem treated we wait for a total of at most r failures, r being specified in advance. A statistic $V_r^{m,n}$ is defined in terms of the ranks of the observations from each population. It is shown to be equivalent to a generalized Wilcoxon statistic and a generalized Mann-Whitney statistic for censored samples; for $r = m + n - 1$ or $r = m + n$ it is equivalent to the usual Wilcoxon test. The test based on $|V_r^{m,n}|$ is studied from the point of view of power, expected number of failures and expected time until termination; numerical comparisons are made with other tests under certain alternatives for small samples. Asymptotic normality for the null case follows from the results of Wald and Wolfowitz [14] and the non-null case is considered by Basu [2].

*Work on this paper was begun while the author was with the Bell Telephone Laboratories, Allentown, Pennsylvania. (see[12]).

2. Introduction.

Let X and Y denote chance variables with continuous unknown distribution functions (c.d.f.) $F = F(x)$ and $G = G(y)$, respectively. A censored sample is observed consisting of the r smallest uncensored and the $N-r$ largest censored observations; m of the N observations have c.d.f. $F(x)$ and are called x's and the remaining $n = N-m$ have c.d.f. $G(y)$ and are called y's. Clearly the number of x's among the r smallest observations is a random variable.

This sample is to be used to test the hypothesis that $P\{X < Y\} = P\{X > Y\}$ against the alternative that these are not equal. The same test is also used to test H_0: $F = G$ against alternatives in which the median is changing monotonically. For example, it would be appropriate if we were dealing with translation alternatives. The corresponding one-sided test can also be used against one-sided alternatives of the form $F(x) > G(x)$ for all x.

One area where this type of problem arises is in life-testing where the observations are the times of failure. The experimenter puts $N = m+n$ units on test, m of one kind and n of another, and he terminates the experiment after r failures so that the censored observations are the unfailed units. We assume that r is preassigned. The test is nonparametric and the test statistic depends only on the order of the x's and y's and the number of each that are among the first r failures.

A statistic $V_r^{m,n} = V_r$ is introduced for this problem in section 3 and its relation to other statistics is considered. In particular it is shown to be equivalent to a generalization of the Wilcoxon [15] and Mann-Whitney [10] statistics. The exact distribution is derived in section 4. In section 5 we discuss the curtailed form of the two tail test based on $|V_r|$. Formulas for the power of the test, the expected number of failures and the expected time required by the test are derived for two classes of alternatives in section 6. In section 7 these characteristics are used to make small sample comparisons between our statistic and several others that have been proposed. Asymptotic normality of V_r is shown in section 8.

Asymptotic properties of the test procedure based on V_r, including the consistency and the asymptotic distribution under the alternative $F \neq G$ have been recently considered by Basu [2].

Alling [1] considers a curtailed test in which failures are observed only until the decision based on the "complete" Wilcoxon statistic is determined; this is equivalent to our test based on $|V_r|$ for $r = N$ and $r = N-1$.

A related problem is considered by Halperin [9] where the test terminates at a preassigned time and r becomes a chance variable. More recently Gehan [8] has extended Halperin's formulation to include the case in which units are put on test at different times so that different units have different censoring points even though the experiment is stopped at a fixed time; tied observations are also allowed in this formulation. The Halperin-Gehan statistic, as it applies to our problem is included in the comparisons of section 7.

This paper is also related to the work of Rao, Savage and Sobel [11] which considers locally most powerful (LMP) censored tests against various alternatives. The LMP test against the Lehman alternatives, taken from [11], is included in the comparisons in section 7.

3. Definitions of V_r and W_r.

In this section we define our statistic $V_r = V_r^{m,n}$ based on r observed failures out of the combined set of $N = m+n$ units, m units (called x's) with c.d.f. $F(x)$ and n units (called y's) with c.d.f. $G(y)$. We also define a related statistic $W_r = W_r^{m,n}$ which helps to motivate V_r.

Let m_i and n_i denote the cumulative number of failures from $F = F(x)$ and $G = G(y)$, respectively, up to and including the i^{th} failure so that for each i

$$(3.1) \qquad m_i + n_i = i \qquad (i = 1,2,\ldots,r).$$

Let E_0 denote expectation under the null hypothesis H_0 that $F = G$. For any ordered sequence of x's and y's of length r, let

$$(3.2) \qquad \begin{aligned} v_i &= nm_i - mn_i = mn\left(\frac{m_i}{m} - \frac{n_i}{n}\right) && (i = 1,2,\ldots,r) \\ v_j &= nE_0\{m_j|m_r\} - mE_0\{n_j|n_r\} = n\left[m_r+(j-r)\left(\frac{m-m_r}{N-r}\right)\right] - m\left[n_r+(j-r)\left(\frac{n-n_r}{N-r}\right)\right] \\ &= (nm_r-mn_r)\left(\frac{N-j}{N-r}\right) && (j = r+1,r+2,\ldots,N). \end{aligned}$$

The proposed statistic (for any such sequence) is defined by

$$(3.3) \qquad V_r = \sum_{i=1}^{N} v_i = \sum_{i=1}^{r}(nm_i-mn_i) + \left(\frac{N-1-r}{2}\right)(nm_r-mn_r)$$

and we shall denote the uncensored portion of V_r, i.e., the sum on the right side of (3.3), by V_r'. We note that $V_N = V_{N-1} = V_N' = V_{N-1}'$ and for any r, both V_r and V_r' depend on the order of the x's and y's in the sequence.

One possible motivation of the statistic is that under H_0 we expect to have m_i/m equal to n_i/n for each i and hence to have $V_r = 0$. For example, if $m = 3$, $n = 2$, $r = 4$ and we observe xyxy then $u_1 = 2$, $u_2 = -1$, $u_3 = 1$, $u_4 = -2$, $u_5 = 0$ and hence $V_4 = 0$. Other motivations for our statistic will appear later; for example, for $r = N-1$ (and $r = N$) it reduces to the Wilcoxon statistic. The statistic V_r' was introduced by the author in [12].

Following the same line of argument as given in Basu [2] it can be shown that the asymptotic distribution of V_r under the alternatives is again normal.

An alternative definition of V_r is based on the reverse ranks of the r failures, i.e., counting the r^{th} failure as 1, the $r-1^{st}$ as 2, etc. Define $\delta_x(i)$ to be i if there is an x in the i^{th} position and zero otherwise; define $\delta_x^*(i)$ to be i if there is an x in the $(r+1-i)^{th}$ position and zero otherwise $(i = 1,2,\ldots,r)$. Define $\delta_y(i)$ and $\delta_y^*(i)$ similarly. Then $\delta_x(i) + \delta_y(i) = i$, $\delta_x^*(i) + \delta_y^*(i) = i$ and

$$(3.4) \qquad \sum_{i=1}^{r} \delta_x(i) + \sum_{i=1}^{r} \delta_x^*(i) = (r+1)m_r; \quad \sum_{i=1}^{r} \delta_y(i) + \sum_{i=1}^{r} \delta_y^*(i) = (r+1)n_r .$$

The statistic W_r is defined as the sum of the scores of the x's, where each uncensored x has its rank as a score and each censored x is scored as $(N+r+1)/2$, the average of the numbers from $r+1$ to N inclusive. We wish to show that the statistics V_r and W_r are equivalent for any r. Let W_r' denote the uncensored portion of W_r, i.e., the first sum in (3.4), so that

$$(3.5) \qquad W_r = W_r' + (m-m_r)\left(\frac{N+r+1}{2}\right) .$$

As an auxiliary statistic we also define

$$(3.6) \qquad D_r' = \frac{N}{2} \left[\sum_{i=1}^{r} \delta_x^*(i) - \sum_{i=1}^{r} \delta_y^*(i) \right] ,$$

and we note that (with or without stars) the 2 sums in (3.6) add to the constant $\binom{r+1}{2}$.

To show the equivalence we first prove

<u>Lemma 1</u>: For any r

$$(3.7) \qquad \sum_{i=1}^{r} \delta_x^*(i) = \sum_{i=1}^{r} m_i \quad ; \quad \Sigma\delta_y^*(i) = \sum_{i=1}^{r} n_i .$$

Proof: Consider the j^{th} position from the point of censoring. If it is an x then it contributes j to the first sum in (4.1) and it contributes 1 to each of $m_{r+1-j}, m_{r+2-j}, \ldots, m_r$, i.e., it adds j to Σm_i. If there is a y in the j^{th} position it contributes zero to both sides of the equation. This proves the first equation in (3.7); the proof of the second is similar and is omitted.

Lemma 2: The statistics V_r' and D_r' are equivalent.

Proof: Using (3.1) and the definition of V_r' after (3.3) we obtain

$$V_r' = N \sum_{i=1}^{r} m_i - m\binom{r+1}{2} = - N \sum_{i=1}^{r} n_i + n\binom{r+1}{2}. \tag{3.8}$$

Taking ½ the sum of these 2 results and using lemma 1 gives the desired result

$$V_r' = D_r' + \left(\frac{n-m}{2}\right)\binom{r+1}{2} . \tag{3.9}$$

Lemma 3: The statistics V_r and W_r are equivalent.

Proof: Using (3.4) and (3.5) we can write D_r' in terms of the "forward" ranks $\delta_x(i)$ as

$$D_r' = \frac{N}{2}\left[m(N+r+1) - \binom{r+1}{2} - 2W_r - m_r(N-r-1)\right].$$

Substituting this and the definition of V_r' from (3.3) into (3.9) gives the desired result

$$V_r = m\binom{N+1}{2} - NW_r . \tag{3.10}$$

For $r = N$ we note that V_r is equivalent to the uncensored Wilcoxon or Mann-Whitney statistic U (defined in[10]). Let U_r' denote the number of pairs (x,y) with $y < x$ among the r uncensored observations. To this we add $n_r(m-m_r)$ for the pairs (x,y) with x censored and y uncensored and also the expected number of such pairs, $\frac{1}{2}(m-m_r)(n-n_r)$, under H_0 for the pairs (x,y) with both x and y censored, and we define the result to be U_r, i.e.,

$$U_r = U_r' + \frac{(m-m_r)(n+n_r)}{2} . \tag{3.11}$$

Lemma 4: The statistics U_r and V_r are equivalent.

Proof: Starting with (3.10), we replace W_r by W_r' using (3.5). For any fixed m_r and n_r we make use of the known relation (with n replaced by r) between the Wilcoxon and Mann-Whitney statistic, viz.

$$W_r' = U_r' + \binom{m_r+1}{2} = \sum_{i=1}^{r} \delta_x(i) . \tag{3.12}$$

Finally, using (3.11), we replace U_r' by U_r obtaining

(3.13) $$V_r = N(\frac{mn}{2} - U_r).$$

We conclude this section by proving

Lemma 5: If $m_r = m$ or $n_r = n$ for some $r = r_0$ then the value of V_r is the same for all $r \geqq r_0$.

Proof: If $m_r = m$ then $m_{r+1} = m$, $n_{r+1} = n_r+1$ and $r-n_r = m$. Hence, for $r \geqq r_0$ from (3.3)

(3.14) $$V_{r+1}-V_r = [(n-1)m - mn_r](\frac{N-r}{2}) - (nm - mn_r)(\frac{N-1-r}{2})$$

$$= - m(\frac{N-r}{2}) + (\frac{nm-mn_r}{2}) = \frac{m}{2}(r-n_r) - \frac{m^2}{2} = 0.$$

The proof for the case $n_r = n$ is similar and is omitted.

This property makes the statistic V_r more desirable than V_r' since it shows that the value of V_r will not change if one waits for further failures when all the units from one source have already failed; the corresponding property does not hold for V_r'.

4. Exact Distribution and Moments of V_r under H_0.

In this section we obtain the exact distribution of V_r under H_0 in terms of the distribution $\pi(u|m,n)$ of the Mann-Whitney statistic U or in terms of a partition function $A(u,m,n)$ defined in [6] as the number of ways it is possible to select exactly m nonnegative integers, none greater than n, whose sum does not exceed u. For any fixed pair m_r, n_r we start with U_r' with c.d.f $\pi(u|m_r,n_r)$ based on r observations defined above. Using (3.10) and (3.12) the symmetrical dual of U_r' defined as $U_r'' = m_r n_r - U_r'$ is given in terms of V_r by

$$(4.1) \qquad U_r'' = m_r n_r - U_r' = \frac{m_r}{2}(2r+1-m_r) + \frac{V_r}{N} - \frac{m(N+1)}{2} = h(V_r) \text{ (say).}$$

By the symmetry of U_r' about $m_r n_r$ it follows that U_r'' has the same distribution as U_r'. Since $h(v)$ is a monotonic function of v, we have from (4.1) and the first page of [6], letting $u = h(v)$, that under H_0

$$(4.2) \qquad P\{V_r \leq v|m_r,n_r\} = \pi(u|m_r,n_r) = A(u,m_r,n_r)/\binom{r}{m_r}.$$

Hence the unconditional distribution of V_r under H_0 is

$$(4.3) \qquad P\{V_r \leq v\} = \sum_{m_r} \frac{A(u,m_r,n_r)}{\binom{r}{m_r}} \frac{\binom{r}{m_r}\binom{n}{r-m_r}}{\binom{N}{r}} = \frac{1}{\binom{N}{m}} \sum_{m_r} A(u,m_r,n_r)\binom{N-r}{m-m_r}$$

where the limits of summation on m_r need only run from Max(0,r-n) to Min(r,m). The function $A(u,m,n)$ is easily computed for small m and n (see [6] for details) and for larger values the tables in [6] are useful.

For m = n and any r we can show that V_r is symmetrical under H_0 by defining for each sequence a complementary sequence (with the same probability under H_0) obtained by interchanging x's and y's. Since $E_0(V_r)$ is shown to be zero below this symmetry must be about zero.

Of course the probability under H_0 of any given sequence S_r of length r with m_r x-components and n_r y-components is given by

$$(4.4) \qquad P_0\{S_r|m_r,n_r\} = \frac{\binom{m}{m_r}\binom{n}{n_r}}{\binom{r}{m_r}\binom{N}{r}} = \frac{\binom{N-r}{m-m_r}}{\binom{N}{m}}.$$

We can also regard the derivation of the distribution of W_r' (and hence also V_r) under H_0 as a finite urn (or card) problem with r balls marked 1 to r and N-r balls all marked $(N+r+1)/2$; then W_r' is the sum of the scores obtained by selecting m balls at random without replacement.

The first 4 moments of V_r under H_0 will be needed below; we now derive them. For $\alpha = 1,2,\ldots,r$ let

(4.5) $$t_\alpha = \begin{cases} n & \text{if } \alpha^{th} \text{ observation is an } x \\ -m & \text{if } \alpha^{th} \text{ observation is a } y. \end{cases}$$

Then $v_i = t_1+t_2+\ldots+t_i$ for $i = 1,2,\ldots,r$ and $v_j = v_r(N-j)/(N-r)$ for $j > r$. Hence by (3.3)

(4.6) $$V_r = \sum_{i=1}^{r} v_i + \sum_{j=r+1}^{N} v_j = \sum_{\alpha=1}^{r} (r+1-\alpha)t_\alpha + (\frac{N-r-1}{2}) \sum_{\alpha=1}^{r} t_\alpha = \sum_{\alpha=1}^{r} (\frac{N+r+1}{2} - \alpha)t_\alpha .$$

Clearly $E_0\{t_\alpha\} = 0$ for all m, n, α and hence by (4.6) and (3.3)

(4.7) $$E_0\{V_r\} = E_0\{V_r'\} = 0.$$

For the covariance of any two t's under H_0 we easily obtain

(4.8) $$\sigma_0(t_\alpha, t_\beta) = \begin{cases} -\frac{mn}{N-1} & \text{for } \alpha \neq \beta;\ \alpha \leq r,\ \beta \leq r \\ mn & \text{for } \alpha = \beta \leq r \end{cases}$$

and hence for $i \leq j$ as an auxiliary result we have

(4.9) $$\sigma_0(v_i, v_j) = \begin{cases} \frac{mni(N-j)}{N-1} & \text{for } i \leq r \\ \frac{mnr(N-i)(N-j)}{(N-1)(N-r)} & \text{for } r < i. \end{cases}$$

From (4.6) and (4.8) (or from (4.6) and (4.9)) we obtain after simplification

(4.10) $$\sigma_0^2 (V_r) = \frac{mnrN}{12(N-1)} \{3N(N-r)+r^2-1\}.$$

The third and fourth moments of V_r under H_0 are similarly obtained; the final results for $N > 2$ and $N > 3$, respectively, are

(4.11) $$E_0\{V_r^3\} = \frac{mn(n-m)Nr(N-r-1)(N-r)(N-r+1)}{8(N-1)(N-2)} ,$$

$$(4.12) \qquad E_0\{V_r^4\} = \frac{mnr}{240}\{BT_1 + \frac{mnNT_2 - BT_3}{N-1} + \frac{2B-mnN}{(N-1)(N-2)}(T_4 - \frac{T_5}{N-3})\}$$

where $B = m^2 + n^2 - mn$ and the T_i's are given by

$$T_1 = 15N^4 + 30N^2(r^2-1) + (r^2-1)(3r^2-7),$$

$$T_2 = (r-1)[45N^4 + 30N^2(r+1)(r-3) + (r+1)(5r^3 - 9r^2 - 5r + 21)],$$

$$T_3 = (r-1)[105N^4 + 30N^2(r+1)(3r-7) + (r+1)(5r^3 - 21r^2 - 5r + 49)],$$

$$T_4 = 2(r-1)(r-2)[45N^4 + 15N^2(r+1)(r-6) - (r+1)(r-3)(5r+7)],$$

$$T_5 = 3(r-1)(r-2)(r-3)[15N^4 - 30N^2(r+1) + (r+1)(5r+7)].$$

We note that the third (central) moment vanishes for $m = n$ (any r) and for $r = N$ and $r = N-1$ (any m,n). If $N = 2$ or 3 then (4.11) and (4.12) still give the correct result if an equal number of zero factors in the numerator and denominator are cancelled.

Exact probabilities of V_r under H_0 for $m = n = r = 4(1)8$ are given in Table I. The integers in the second column have to be divided by a common denominator D (given at the head of the column) to obtain the required probability. Thus the second entry for $m = n = r = 6$ shows that $6/924 = .0065$ is the probability under H_0 that $V_6 = 174$.

Table I: Distribution of $V_n^{n,n} = V_n$ for $n = 4(1)8$

V_4	Indiv. (D=70)	Cumulative
64	1	.01429
44	4	.07143
36	4	.12859
28	4	.18571
20	4	.24286
16	6	.32857
8	6	.41429
0	12	.58571

V_7	Indiv. (D=3432	Cumulative
343	1	.00029
287	7	.00233
273	7	.00437
259	7	.00641
245	7	.00845
231	7	.01049
217	28	.01865
203	28	.02681
189	42	.03904
175	42	.05128
161	63	.06964
147	63	.08800
133	98	.11655
119	77	.13899
105	112	.17162
91	126	.20833
77	161	.25524
63	140	.29604
49	175	.34703
35	175	.39802
21	175	.44901
7	175	.50000

V_5	Indiv. (D=252)	Cumulative
125	1	.00395
95	5	.02381
85	5	.04365
75	5	.06349
65	5	.08333
55	15	.14286
45	10	.18254
35	20	.26190
25	20	.34127
15	20	.42063
5	20	.50000

V_6	Indiv. (D=924)	Cumulative
216	1	.00108
174	6	.00758
162	6	.01407
150	6	.02056
138	6	.02706
126	6	.03355
120	15	.04978
114	6	.05628
108	15	.07251
96	30	.10498
84	30	.13745
72	45	.18615
60	30	.21861
54	20	.24026
48	30	.27273
42	20	.29437
36	15	.31061
30	40	.35390
24	15	.37013
18	60	.43506
6	60	.50000

V_8	Indiv. (D=12870)	Cumulative
512	1	.00008
440	8	.00070
424	8	.00132
408	8	.00194
392	8	.00256
376	8	.00319
360	8	.00381
352	28	.00598
344	8	.00660
336	28	.00878
328	8	.00940
320	56	.01375
304	56	.01810
288	84	.02463
272	84	.03116
256	112	.03986
248	56	.04421
240	84	.05074
232	56	.05509
224	84	.06162
216	112	.07032
200	168	.08337
208	56	.08772
192	56	.09207
184	224	.10948
176	28	.11166
168	280	.13341
160	28	.13559
152	336	.16169
136	336	.18780
128	70	.19324
120	336	.21935
112	70	.22479
104	336	.25089
96	140	.26177
88	280	.28353
80	210	.29984
72	224	.31725
64	350	.34444
56	168	.35750
48	350	.38469
40	112	.39339
32	490	.43147
24	56	.43582
16	490	.47390
8	56	.47825
0	560	.52176

5. The Test Based on $|V_r|$ and its Curtailed Form.

In this section we construct the test based on $|V_r|$ with size α for testing the hypothesis that $P\{X > Y\} = P\{Y > X\}$ against the alternative that these probabilities are not equal. We then use the same test to test H_0: $F \equiv G$ against any alternatives in which the median is changing monotonically. For example, this would be appropriate if we were dealing with a family which is generated by varying a single location parameter.

Consider the case $m = n = r = 6$. Since $m = n$, V_6 is symmetric under H_0 and we use an equal tail test based on $|V_6|$. Sixteen sequences with the largest values of $|V_6|$ are shown in Table II; only eight rows are needed because of the duality that is present; the sequence S^* is dual to S if it is obtained from S by interchanging x's and y's.

Table II: Test Based on $|V_6|$ for $m = n = 6$

Sequence S	Dual Sequence S^*	$\|V_6\|$	$P_0(S) + P_0(S^*) = 2P_0(S)$ Indiv.	Cumulative
xxxxxx	yyyyyy	216	1/462	.0022
xxxxxy	yyyyyx	174	6/462	.0152
xxxxyx	yyyyxy	162	6/462	.0281
xxxyxx	yyyxyy	150	6/462	.0411
xxyxxx	yyxyyy	138	6/462	.0541
xyxxxx	yxyyyy	126	6/462	.0671
xxxxyy	yyyyxx	120	15/462	.0996
yxxxxx	xyyyyy	114	6/462	.1126
⋮	⋮	⋮	⋮	⋮

The proposed test is to reject H_0 for large values of $|V_6|$. To obtain an α of exactly .05 we reject H_0 when $|V_6| > 138$, accept (or fail to reject) H_0 when $|V_6| < 138$ and randomize when $|V_6|$ equals the critical value $|V_r|^c = 138$. More precisely, if $|V_6| = 138$ we perform an independent experiment which will reject H_0 with probability .68.

If randomization is not allowed then we might still consider putting one of the two sequences with $|V_6| = 138$ in the rejection region (and the other in the acceptance region) but in this case it would destroy the symmetry of the test and we do not consider it.

It is evident that the result of the test may be determined before 6 failures are observed and hence the test can be put in a curtailed form. Table III gives the results for the above example allowing randomization. We use the symbols $|V_r|_R$ and $|V_r|_{\bar{R}}$ according as randomization is or is not allowed. Since the test is symmetric we can restrict the tabulation to those sequences starting with an x. Let $E_0\{N_f\}$ denote the expected number of failures required by the test under H_0; for the example in Table III we obtain $E_0\{N_f\} = 3.348$.

Table III: Test Based on $|V_6^{6,6}|_R$ in Curtailed Form

Stopping Sequences S	$2P_0(S)$	$\lvert V_6\rvert$	Action
xxxxx	7/462		
xxxxyx	6/462	$\lvert V_6\rvert > 138$	Reject H_0
xxxyxx	6/462		
xxyxxx	6/462	$\lvert V_6\rvert = 138$; $P\{\text{Reject } H_0\} = .68$	
xxxxyy	15/462		
xxxyxy	15/462		
xxxyy	35/462		
xxyxxy	15/462	$\lvert V_6\rvert < 138$	Accept H_0
xxyxy	35/462		
xxyy	70/462		
xy	252/462		

It should be pointed out that the caluculation of Table III, which may be tedious for large values of m, n and r, is not necessary for carrying out the test. The critical value $|V_r|^c$ can be obtained by means of the normal approximation when r is not too small (see discussion in section 8).

Briefly the curtailed test is to stop as soon as the decision (or action) to be taken is determined. In the above example suppose we observe xxy initially. Then we compute the smallest and largest values that V_6 can attain with 3 more observations; these are -24 and 138. Since they do not all lead to the same action with probability one, we wait for another failure. If we then obtain xxyy then the possible extremes are -24 and 72. All possible values now lead to the same action and we terminate the test (accepting H_0).

Another interesting property of the statistic V_r is concerned with the result obtained by using (3.3) or (4.6) with r replaced by d for any curtailed sequence S of length $d < r$.

Lemma 6: For any curtailed sequence of length $d \leq r$ the value V_d obtained by using (3.3) with r replaced by d is the conditional expectation under H_0 of V_r given the source of the first d failures.

Proof: From (3.3) we have

$$(5.1) \qquad E_d\{V_r\} = \sum_{i=1}^{d} (nm_i - mn_i) + E_d\{\sum_{i=d+1}^{r} (nm_i - mn_i)\} + \left(\frac{N-r-1}{2}\right) E_d\{m_r - mn_r\},$$

where E_d denotes the conditional expectation given the first d failures. For $i > d$, using the hypergeometric distribution, we obtain

$$(5.2) \quad E_d\{m_i\} = m_d + (i-d)\left(\frac{m-m_d}{N-d}\right) \quad ; \quad E_d\{n_i\} = n_d + (i-d)\left(\frac{n-n_d}{N-d}\right).$$

Substituting these in (5.1) and letting $\Delta = nm_d - mn_d$ gives

$$(5.3) \qquad E_d\{V_r\} = \sum_{i=1}^{d} (nm_i - mn_i) + \Delta\left[r-d- \frac{\binom{r-d+1}{2}}{N-d} + \left(\frac{N-r-1}{2}\right)\left(1- \frac{r-d}{N-d}\right)\right]$$

$$= \sum_{i=1}^{d} (nm_i - mn_i) + \left(\frac{N-d-1}{2}\right)\Delta = V_d,$$

which is the desired result.

Lemma 5 can be regarded as a special case of the above lemma in which the random variable is constant if the source of the first d failures is given. Lemma 6 proves that the observed values of V_r for increasing r form a martingale.

6. Formulas for the Power and Expected Time Under Two Alternatives.

In this section we derive formulas for the power of the test based on $|V_r|$ (for m = n we shall write $V_r^{(n)}$ or V_r^n when there is no danger of confusion) for two classes of alternatives; numerical computations are carried out for one particular alternative in each class. These formulas can be used for any nonparametric test.

Epstein [5] has made some sampling (i.e., Monte Carlo) studies of the power of several nonparametric tests. We wish to compare exact results for the test based on $|V_r|$ for $\alpha = .05$ and a small common value of m = n with the tests he considers. Epstein considered a run test, a rank sum test (not the same as W_r above), a set of exceedance tests based on E_r^n (r = 1, 2, 3; n = 10) and a set of maximum deviation tests based on M_r^n (r = 1, 3, 6, 10; n = 10); E_r^n was studied by Epstein [4] and M_r^n by Tsao [13]. A statistic $\hat{U}_r'^{m,n}$ which scores +1 for each pair (x,y) with y < x and -1 for pairs with x < y (and zero if x and y are both censored) was considered by Halperin [9] and Gehan [8]; we include a test based on $|\hat{U}_r'^{m,n}|$ in our small sample comparisons and denote it by $|\hat{U}_r'^{m}|$ for m = n. Another statistic P_r^n called the precedence life test statistic was studied by Eilbatt and Nadler [3].

We wish to compare some of these tests with a corresponding test based on $|V_r^n|$ with the same n. We compare not only the power P{Correct Decision $|H_i$ } = P_i(CD} under alternatives H_i (i = 1, 2) but also the expected number of failures $E_i\{N_f\}$ under H_i and the expected time required by the curtailed test $E_i\{T\}$ under H_i (i = 0, 1, 2). The rest of this section is devoted to describing the two particular alternatives selected and the derivation of special formulas for these alternatives. A general formula for the probability of any rank order under any alternative appears in Rao, Savage, Sobel [11] and our $P_i\{CD\}$-expressions can be regarded as special cases of this.

We consider two sets of alternatives denoted as $H_1^{(p)}$ and $H_2^{(c)}$. Under $H_1^{(p)}$ the two cumulative distribution functions F(x) and G(y) have the respective densities (omitting the values of x and y where the density is zero)

(6.1) $$f_\theta(x) = \frac{1}{\theta} e^{-x/\theta} \qquad x \geq 0,$$

(6.2) $$g_\theta(y) = \frac{1}{\theta(1-p)} e^{-y/\theta} \qquad y \geq \theta \ln(\frac{1}{1-p}) .$$

In the numerical calculations of Table IV only the case $p = \frac{1}{2}$, denoted by H_1, is considered. Under $H_2^{(c)}$ the density of $f_\theta(x)$ is as in (6.1) and

(6.3) $$g_\theta(y) = \frac{1}{c\theta} e^{-y/c\theta} \qquad y \geq 0,$$

so that one has a mean that is c times the mean of the other. This is a so-called Lehmann alternative since for any t we have $[1-G(t)]^c = 1-F(t)$. In the numerical calculations of Table IV only the case $c = 2$, denoted by H_2, is considered. These two alternatives were clearly chosen because of their interest in life testing applications.

In order to compute the power it is necessary to first develop some formulas for the probability of observing a particular sequence of x's and y's under $H_1^{(p)}$ and $H_2^{(c)}$. Let $X_1, X_2, \ldots, X_i$ and $Y_1, Y_2, \ldots, Y_j$ denote the ordered X's and Y's in a sample of size $d = i + j$ where $i \leq m$, $j \leq n$ and $d \leq r$. Let R_i denote the ranks of Y_i in the combined sequence S_d of length d $(i = 1,2,\ldots,j)$ and let $P_i\{S_d\}$ denote the probability of S_d under H_i $(i = 0,1,2)$.

Case 1: Suppose $j = 0$ so that $d \leq m$ and we observe only x's. Then

(6.4) $$P_1\{S_d\} = \frac{m!}{(d-1)!(m-d)!} \int_0^{\theta\ln(\frac{1}{1-p})} F(x_d)^{d-1} [1-F(x_d)]^{m-d} dF(x_d)$$

$$+ \frac{m!}{(d-1)!(m-d)!} \int_{\theta\ln(\frac{1}{1-p})}^{\infty} F(x_d)^{d-1} [1-F(x_d)]^{m-d} [1-G(x_d)]^n dF(x_d)$$

$$= I_p(d,m-d+1) + \frac{\binom{N-d}{m-d}}{(1-p)^n \binom{N}{m}} I_{1-p}(N+1-d,d)$$

where $I_p(a,b)$ is the standard notation for the incomplete beta function; here we have used the fact that for $t \geq \theta \ln[1/(1-p)]$

$$(6.5) \qquad 1-G(t) = \frac{1-F(t)}{1-p}$$

and made the transformation $u = F(x_d)$.

Case 2: Suppose $j > 0$ and $R_j = d$ so that the d^{th} observation is y_j. Then

$$(6.6) \qquad P_1\{S_d\} = \frac{i!j!\binom{m}{i}\binom{n}{j}}{\prod_{\alpha=1}^{j}(R_\alpha - R_{\alpha-1}-1)!} \int_{\theta\ln(\frac{1}{1-p}) < y_1 < y_2 < \ldots < y_j < \infty} \cdots \int [1-G(y_j)]^{n-j}\,[1-F(y_j)]^{m-i} \cdot$$

$$\prod_{\alpha=1}^{j} [F(y_\alpha)-F(y_{\alpha-1})]^{R_\alpha - R_{\alpha-1}-1}\, dG(y_\alpha)$$

$$= \frac{\binom{N-d}{m-i}}{\binom{N}{m}} (1-p)^{-n}\, I_{1-p}(N+1-R_1, R_1) .$$

To obtain the above we used (6.5) and we iteratively integrated out y_j, y_{j-1}, etc., leaving only the integral on y_1; here (and in the next case below) $y_0 = 0$ and $R_0 = 0$. The details are straightforward and are omitted.

Case 3: Suppose $j > 0$ and $R_j < d$ so that the d^{th} observation is $x_i = x$ (say). Then

$$(6.7) \qquad P_1\{S_d\} = \frac{i!j!\binom{m}{i}\binom{n}{j}}{(d-R_j-1)!\prod_{\alpha=1}^{j}(R_\alpha - R_{\alpha-1}-1)!} \int_{\theta\ln(\frac{1}{1-p}) < y_1 < \ldots < y_j < x < \infty} \cdots \int [1-G(x)^{n-j}\,[1-F(x)]^{m-i} \cdot$$

$$[F(x)-F(y_j)]^{d-R_j-1} \left\{ \prod_{\alpha=1}^{j} [F(y_\alpha)-F(y_{\alpha-1})]^{R_\alpha - R_{\alpha-1}-1}\, dG(y_\alpha) \right\} dF(x)$$

$$= \frac{\binom{N-d}{m-i}}{(1-p)^n \binom{N}{m}}\, I_{1-p}(N+1-R_1, R_1) .$$

This result is the same as in (6.6); the derivation is similar to that in Case 2 and is omitted. Thus we note from (6.4), (6.6) and (6.7) that in all cases the $P\{S_d|H_1\}$ depends only on R_1, the rank of the first y in the combined sequence.

For p = 0 all three give the same result for H_0, namely,

$$(6.8)\qquad P_0\{S_d\} = \binom{N-d}{m-i}/\binom{N}{m}$$

which agrees with (4.4) for d = r.

To compute the expected time $E\{T|H_1\}$ under H_1 we again consider the individual terms of the result corresponding to any particular stopping sequence S_d. We consider the same three cases as above; the resulting expressions again depend only on d and R_1 and again Cases 2 and 3 give the same result. The results depend on a lemma dealing with a function J(x,y) defined for x > 0, y > 0 by

$$(6.9)\qquad J_q(x,y) = \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \int_{0+}^{q} \ln(\tfrac{1}{u})\, u^{x-1}(1-u)^{y-1}\, du;$$

for convenience we define $J_q(x,0)$ to be zero for any x > 0 and any $q(0 < q \leq 1)$.

Lemma. For x > 0, y > 1 and $0 \leq q \leq 1$

$$(6.10)\qquad J_q(x,y) = \frac{[\ln(\frac{1}{q})]\Gamma(x+y)}{\Gamma(x+1)\Gamma(y)}\, q^x(1-q)^{y-1} + \frac{I_q(x,y)}{x} + J_q(x+1, y-1).$$

If $y \geq 1$ is an integer and $q \leq 1$ then we can iterate (6.10) and letting s = x+y we obtain

$$(6.11)\qquad J_q(s-y, y) = \ln(\tfrac{1}{q})\, I_q(s-y, y) + \sum_{\alpha=1}^{y} \frac{I_q(s-\alpha, \alpha)}{s-\alpha}\,.$$

In particular, for q = 1 we obtain from (6.11) a sum of reciprocals and it is easily verified that for any integers r > s > t > u

$$(6.12)\qquad J_1(r-s, s-t) + J_1(r-t, t-u) = J_1(r-s, s-u)\,.$$

This lemma can be obtained by starting with a simple integration by parts in (6.9) to obtain (6.10). Then (6.11) is obtained by iteration and (6.12) is a consequence of (6.11) for q = 1. The details are omitted.

It should also be noted that we can write

$$(6.13)\qquad J_1(x,y) = \sum_{j=1}^{x+y-1} \frac{1}{j} - \sum_{j=1}^{x-1} \frac{1}{j} = D(x+y-1) - D(x-1)$$

where $D(x)$ is the well-tabulated digamma function; we note that the first and largest fraction in $J(x,y)$ is $1/x$ and y is the number of fractions. Consider a stopping sequence $S_d = S_d^{(i,j)}$ of length d, with i x-observations and $j = d-i$ y-observations; suppose S_d falls in Case 2, i.e., it ends in y_j. Let $E_1^*\{T|S_d\}$ denote the contribution to (i.e., the term in) the expected time (under H_1) corresponding to the stopping sequence S_d; these are not expectations but the numerators of conditional expectations. Using the above lemma we obtain for

Case 2: $R_j = d$

$$(6.14) \qquad E_1^*\{T|S_d\} = \frac{\theta\binom{N-d}{m-i}}{(1-p)^n\binom{N}{m}} \left\{ \sum_{\alpha=0}^{R_1-1} \frac{I_{1-p}(N-\alpha,\ \alpha+1)}{N-\alpha} + I_{1-p}(N+1-R_1,\ R_1)\ [\ln(\tfrac{1}{1-p})+J_1(N+1-d,\ d-R_1)] \right\} .$$

Sketch of Proof: To obtain (6.14) we start with

$$(6.15) \qquad E_1^*\{T|S_d\} = \frac{i!\binom{m}{i}j!\binom{n}{j}}{\prod_{\alpha=1}^{j}(R_\alpha - R_{\alpha-1}-1)!} \int_{\theta\ln(\frac{1}{1-p}) < y_1 < y_2 <\dots< y_j < \infty} \cdots \int y_j[1-G(y_j)]^{n-j}[1-F(y_j)]^{m-i} \cdot \prod_{\alpha=1}^{j}[F(y_\alpha)-F(y_{\alpha-1})]^{R_\alpha - R_{\alpha-1}-1}\ dG(y_\alpha) ,$$

use (6.5) to eliminate $[1-G(y_j)]$, make the same substitutions $u_j = 1-F(y_j)$ and $w_j = u_j/u_{j-1}$ with $u_0 = 1$ as for (6.6) and write

$$(6.16) \qquad y_j = \log(\frac{1}{w_j}) + \log(\frac{1}{u_{j-1}}) .$$

This gives rise to two integrals; in the first one we can use (6.11) with q=1 as well as (6.12) and the second one is the same as the original integral with j reduced by one. The details are omitted.

It turns out that if S_d falls in Case 3 the result (obtained by the same method as above) is exactly the same as for Case 2 in (6.13). We note that the results for Cases 2 and 3 depend only on d and R_1 as in (6.6) and (6.7).

If S_d contains only x's then we use a similar method in (6.4) (details are omitted) and obtain for Case 1: $R_1 > d$

$$E_1^*\{T|S_d\} = \theta\left[\sum_{\alpha=0}^{d-1} \frac{I_p(\alpha+1,\ m-\alpha)}{m-\alpha} - \ln\left(\frac{1}{1-p}\right) I_{1-p}(m+1-d,\ d)\right] \tag{6.17}$$

$$+ \frac{\theta\binom{N-d}{m-i}}{(1-p)^n\binom{N}{m}} \left[\ln\left(\frac{1}{1-p}\right) I_{1-p}(N+1-d,\ d) + \sum_{\alpha=0}^{d-1} \frac{I_{1-p}(N-\alpha,\ \alpha+1)}{N-\alpha}\right].$$

For p = 0 both (6.14) and (6.17) reduce to the common result for all stopping sequences S_d under H_0

$$E_0^*\{T|S_d\} = \theta \frac{\binom{N-d}{m-i}}{\binom{N}{m}} \sum_{\alpha=0}^{d-1} \frac{1}{N-\alpha} = \theta \frac{\binom{N-d}{m-i}}{\binom{N}{m}} [D(N) - D(N-d)]. \tag{6.18}$$

For the alternative H_2 the results given in (6.3) on the power $P_2\{CD\}$, the expected number of failures $E_2\{N_f\}$ and the expected time $E_2\{T\}$ required until termination are again obtained by treating each stopping sequence S_d separately. We now give the required formulas; derivations are similar to those for H_1 and are omitted.

Case 1: S_d ends in a y so that $R_j = d$.

$$P_2\{S_d\} = \frac{m!}{(m-i)!} \frac{n!}{(n-j)!} \left(\frac{1}{c}\right)^j \prod_{\alpha=1}^{j} \frac{\Gamma(m+1+\frac{n+j-\alpha}{c} - R_{j+1-\alpha})}{\Gamma(m+1+\frac{n+j-\alpha}{c} - R_{j-\alpha})}. \tag{6.19}$$

Case 2: S_d ends in an x so that $R_j < d$.

$$P_2\{S_d\} = \frac{m!}{(m-i)!} \frac{n!}{(n-j)!} \left(\frac{1}{c}\right)^j \prod_{\alpha=1}^{j+1} \frac{\Gamma(m+1+\frac{n+j+1-\alpha}{c} - R_{j+2-\alpha})}{\Gamma(m+1+\frac{n+j+1-\alpha}{c} - R_{j+1-\alpha})}. \tag{6.20}$$

where $R_{j+1} = d$ and $R_0 = 0$.

To compute the expected time $E_2\{T\}$ we derive the contribution

$\therefore E_2^*\{T|S_d\} = P_2\{S_d\}\, E_2\{T|S_d\}$ from each stopping sequence S_d and the sum of these over all stopping sequences yields the value for $E_2\{T\}$. Considering the same two cases as above we obtain for Case 1

$$(6.21) \qquad E_2^*\{T|S_d\} = \frac{m!}{(m-i)!}\,\frac{n!}{(n-j)!}\,\frac{\theta}{c^j}\ \cdot$$
$$\Gamma(N_{j-1}-R_j) \sum_{\beta=1}^{j} \frac{J_1(N_{\beta-1}-R_\beta,\ R_\beta-R_{\beta-1})}{\Gamma(N_{\beta-1}-R_\beta)} \left\{ \prod_{\alpha=1}^{\beta} \frac{\Gamma(N_{\beta-\alpha}-R_{\beta-\alpha+1})}{\Gamma(N_{\beta-\alpha}-R_{\beta-\alpha})} \right\},$$

and for Case 2 we obtain

$$(6.22) \qquad E_2^*\{T|S_d\} = \frac{m!}{(m-i)!}\,\frac{n!}{(n-j)!}\,\frac{\theta}{c^j}\ \cdot$$
$$\Gamma(N_j-R_{j+1}) \sum_{\beta=1}^{j+1} \frac{J_1(N_{\beta-1}-R_\beta,\ R_\beta-R_{\beta-1})}{\Gamma(N_{\beta-1}-R_\beta)} \left\{ \prod_{\alpha=1}^{\beta} \frac{\Gamma(N_{\beta-\alpha}-R_{\beta-\alpha+1})}{\Gamma(N_{\beta-\alpha}-R_{\beta-\alpha})} \right\};$$

in both cases $N_\beta = m+\beta+1+(n-\beta)/c\,(\beta=0,1,\ldots,j+1)$ and $J_1\,(x,y)$ is given by (6.13) since y is an integer.

TABLE IV: COMPARISON OF NONPARAMETRIC CURTAILED LIFE TESTS

$\alpha = .05$ (unless stated otherwise); m = n = 5 for first 6 tests below and m = n = 6 for the last 7 tests below

Test Based On (1)	Critical Value	(2) P_R	(4) Max N_f	Null Hypothesis H_0: (4) $E_0\{N_f\}$	(4) $E_0\{T\}/\theta$	Alternative H_1: $P_1\{CD\}$	$E_1\{N_f\}$	$E_1\{T\}/\theta$	Alternative H_2: $P_2\{CD\}$	$E_2\{N_f\}$	$E_2\{T\}/\theta$
E_2^5	1	.01500	6	4.722	.60595	.16403	5.199	1.09934	.13013	4.928	1.93001
M_2^5	4	.58889	6	4.683	.59802	.23939	4.959	1.05104	.11623	4.854	1.84942
$\lvert LMP\rvert_6^5$	2.028...	.07500	6	4.722	.60959	.17160	5.199	1.09934	.13076	4.928	1.93001
$\lvert\hat{U}_6^5\rvert$	16	.90000	6	4.198	.52262	.42050	4.558	0.71454	.20114	4.336	1.41826
$\lvert V_6^5\rvert_R$	72.5	.60000	6	4.159	.51468	.44428	4.448	0.82890	.19858	4.270	1.37451
$\lvert V_8^5\rvert_R$	80	.80000	8	4.540	.60198	.40342	5.942	1.29917	.21785	4.811	2.10885
E_2^6	1	.53667	7	4.866	.50116	.21624	5.786	1.04414	.12923	5.192	2.03679
M_2^6	4	.27750	7	4.838	.49683	.30801	5.421	0.97078	.12727	5.108	1.96870
$\lvert\hat{U}_7^6\rvert$	24	.64545	7	4.554	.50007	.35072	5.542	0.99094	.12815	5.040	2.07428
$\lvert V_7^6\rvert_R$	144	.64545	7	4.554	.50007	.35072	5.542	0.99094	.12815	5.040	2.07428
(3) $\lvert V_6^6\rvert_R$	138	.68333	6	3.348	.32756	.32013	5.144	0.93309	.12914	3.633	1.12741
(3) $\lvert LMP\rvert_6^6$	1.540...	.68333	6	3.348	.32756	.32013	5.144	0.93309	.12914	3.633	1.12741
(3) $\lvert\hat{U}_6^6\rvert$	23	.68333	6	3.348	.32756	.32013	5.144	0.93309	.12914	3.633	1.12741

(1) E is an exceedance test from [6]; M is a maximum deviation test from [6]; LMP is a locally most powerful test from [10].

(2) P_R denotes the "randomization probability" to achieve $\alpha = .05$.

(3) $\lvert LMP\rvert_6^6$, $\lvert V_6^6\rvert$ and $\lvert\hat{U}_6^6\rvert$ turn out to be identical; also $\lvert V_7^6\rvert$ is identical with $\lvert\hat{U}_7^6\rvert$.

(4) N_f and T denote the number of failures and the time required to terminate the curtailed test.

7. Discussion of Empirical Results.

The numerical results in Table IV show that the test based on $|V_r^n|$ is superior to the exceedance tests and maximum deviation tests for the cases considered. For m = n = 6 the test based on $|V_6^6|$ turned out to be equivalent to the LMP test from [11] but for m = n = 5 the tests based on $|V_6^5|$ and $|V_8^5|$ are both superior to the corresponding LMP test. The performances of the tests based on the $|V_r^n|$ and the $|U_r^n|$ statistics appear to be approximately the same for both alternatives considered.

Table IV also shows that one should not assume that the performances will improve simply by increasing the value of r with fixed m = n or by increasing the common value of m = n with fixed r. It appears that some values of r are better than others for fixed values of m, n; this has not been investigated. The only criterion used in selecting values of m, n and r in Table IV was to make them large enough to serve as typical illustration and not so large that they could not be handled on a desk calculator.

Another test was suggested by the referee of this paper and he claims that it is suggested by the work of Gart [7], but it is clearly not the control median test described in that paper. In the suggested test we take r and N = m+n as above and form the 2×2 table and base the test on the

Population Source	Number of failures before the r^{th}	Number of Censored Observations	Total
X			m*
Y	z		n*
	r-1	N-r	N-1

observed z and the r^{th} observation or (using an approximation) on the associated chi-square χ_1^2) statistic with one degree of freedom

$$(7.1) \qquad \chi_1^2 = \frac{\left(\left|z - (r-1)\frac{n^*}{N-1}\right| - \frac{1}{2}\right)^2}{(r-1)\left(\frac{n^*}{N-1}\right)\left(\frac{m^*}{N-1}\right)\left(\frac{N-r}{N-2}\right)} ,$$

which also contains a so-called continuity correction. Here $m^* = m - 1$, $n^* = n$ if the r^{th} failure is an x and $m^* = m$, $n^* = n - 1$ if the r^{th} failure is a y. Clearly a two-sided test on z-values corresponds to a one-sided test on χ_1^2-values.

This test gives strong weight to the number of failures that are x's and y's and little weight to the order of the failures. For example, for $m = n = 5$, $r = 6$ the 2 sequences xxxyyy and xyxyxy are treated alike; also xxxxyx and yxxxxx are indistinguishable if z(or the associated χ_1^2) is used. Finally it was noted that for the two cases considered in Table IV the exceedence tests (E_2^5 and E_2^6 respectively) were identical to the tests based on z (or the associated χ_1^2) and the tests based on (7.1) were therefore omitted from Table IV.

which also contains a so-called continuity correction. Here $m^* = m - 1$, $n^* = n$ if the r^{th} failure is an x and $m^* = m$, $n^* = n - 1$ if the r^{th} failure is a y. Clearly a two-sided test on z-values corresponds to a one-sided test on χ_1^2-values.

This test gives strong weight to the number of failures that are x's and y's and little weight to the order of the failures. For example, for $m = n = 5$, $r = 6$ the 2 sequences xxxyyy and xyxyxy are treated alike; also xxxxyx and yxxxxx are indistinguishable if z (or the associated χ_1^2) is used.

Finally it was noted that for the two cases considered in Table IV the exceedance tests (E_5^5 and E_5^6 respectively) were identical to the tests based on z (or the associated χ_1^2) and the tests based on (7.1) were therefore omitted from Table IV.

8. Asymptotic Normality of V_r.

We wish to show that under H_0 the distribution of $V_r/\sigma_0(V_r)$ tends to a standard normal distribution as m, n and r all approach infinity so that the triple ratio approaches a fixed triple ratio with positive finite components.

For t_α in (4.6) we can write $nz_\alpha - m(1-z_\alpha) = Nz_\alpha - m$ where $z_\alpha = 1$ or 0 according as the α^{th} observation is an x or a y. Hence we obtain from (4.6)

$$\frac{V_r + \frac{Nrm}{2}}{N^2} = \sum_{\alpha=1}^{r} \left(\frac{N+r+1-2\alpha}{2N}\right) z_\alpha = \sum_{i=1}^{N} a_i z_i \tag{8.1}$$

where for $i = 1,2,\ldots, r$

$$a_i = \begin{cases} \frac{N+r+1-2i}{2N} & \text{if the } i^{th} \text{ observation is an x} \\ 0 & \text{otherwise} \end{cases} \tag{8.2}$$

and $a_i = 0$ for $i = r+1, r+2,\ldots, N$. It is easy to check the conditions of the theorem of Wald and Wolfowitz [14], i.e., to show that

$$\frac{\frac{1}{N}\sum_{i=1}^{N}(a_i - \bar{a}_N)^s}{\left[\frac{1}{N}\sum_{i=1}^{N}(a_i - \bar{a}_N)^2\right]^{s/2}} = \mathcal{O}(1) \quad (s = 3, 4, \ldots) \tag{8.3}$$

and that

$$\frac{\frac{1}{N}\sum_{i=1}^{N}(z_i - \bar{z}_N)^s}{\left[\frac{1}{N}\sum_{i=1}^{N}(z_i - \bar{z}_N)^2\right]^{s/2}} = \mathcal{O}(1) \quad (s = 2, 3, \ldots) \tag{8.4}$$

where $\bar{a}_N = \sum_{i=1}^{N} a_i/N$ and $\bar{z}_N = \sum_{i=1}^{N} z_i/N$; the details are omitted. It follows that the left side of (8.1) and hence V_r is asymptotically normal.

We conclude this section with some empirical remarks about the rapidity of approach to normality. We find that for fairly small values of $m = n$ (with r not too small) we can find the correct critical value $|V_r|^c$ and carry out the test based on $|V_r^n|$ without constructing the tables of stopping sequences as in Tables II and III.

The first approximation of $|V_r|^c$ for the 2-sided test of size $\alpha = .05$ is obtained by computing the closest V_r-values to $\pm$ 1.96 $\sigma_0(V_r)$ where $\sigma_0^2(V_r)$ is given by (4.10). Successive values of $|V_r|$ differ by multiples of N/2 and it is possible to make the usual "continuity-correction" to approximate the probability of particular values of V_r. A useful "rule-of-thumb" (empirical in origin) is to use this one-term normal approximation (NA) when $m \geq 5$, $n \geq 5$ and $r \geq |m-n| + 1/\sqrt{\alpha}$. If this is not satisfied then it is desirable to use exact computations or to use more than one term of the Edgeworth expansion (EA) with continuity-correction

(8.5)
$$P_0\{V_r \leq v\} \approx \Phi(x) - \{\frac{1}{3!}\frac{\mu_3}{\sigma_0^3}\Phi^{(3)}(x)\}$$
$$+ \{\frac{1}{4!}(\frac{\mu_4}{\sigma_0^4} - 3)\Phi^{(4)}(x) + \frac{10}{6!}(\frac{\mu_3}{\sigma_0^3})^2\Phi^{(6)}(x)\} + \ldots$$

where $\Phi(x)$ is the standard normal c.d.f., $\Phi^{(i)}(x)$ is the i^{th} derivative of $\Phi(x)$, μ_i is the i^{th} moment of V_r under H_0 given in section 4, $\sigma_0 = \sqrt{\mu_2}$, $x = v + c'$ and c' (some multiple of N/4) is the continuity correction which depends on the difference of successive values of V_r at the point of interest.

The asymptotic distribution of V_r under the alternatives has been considered by Basu [2].

9. Acknowledgement.

The author wishes to thank Dr. A. P. Basu and the referee of this paper for many useful suggestions. While at the Bell Telephone Laboratories, the author received verbal and written comments on an earlier version of this paper from a number of people, who should also be thanked. Thanks are also due to Miss Marilyn J. Huyett of BTL and Mr. S. P. Yen, Mr. W. J. Park, Mr. Y. L. Tong and Mrs. Maya Weil, all of the University of Minnesota, for assistance with the calculations.

REFERENCES

[1] Alling, D. (1963). Early decision in the Wilcoxon two-sample test. J. Amer. Statist. Assoc., 58, 713-720.

[2] Basu, A. P. (1966). On some two-sample and k-sample rank tests with applications to life testing. (Submitted for publication).

[3] Eilbott, Joan and Nadler, Jack (1956). On precedence life testing. Technometrics, 7, 359-377.

[4] Epstein, B. (1954). Tables for the distribution of the number of exceedances. Ann. Math. Statist., 25, 762-768.

[5] Epstein, B. (1955). Comparison of some non-parametric tests against normal alternatives with an application to life-testing. J. Amer. Statist. Assoc., 50, 894-900.

[6] Fix, Evelyn and Hodges, J. F. (1955). Significant probabilities of the Wilcoxon test. Ann. Math. Statist., 26, 301-312.

[7] Gart, John J. (1963). A median test with sequential application. Biometrika, 50, 55-62.

[8] Gehan, Edmund A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika, 52, 203-223.

[9] Halperin, M. (1960). Extension of the Wilcoxon-Mann-Whitney test to samples censored at the same fixed point. J. Amer. Statist. Assoc., 55, 125-138.

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

[11] Rao, U.V.R., Savage, I. R. and Sobel, M. (1960). Contributions to the theory of rank order statistics: The two-sample case. Ann. of Math. Stat., 31, 415-426.

[12] Sobel, Milton (1957). On a generalized Wilcoxon statistic for life testing. Proc. Working Conference on the theory of reliability (April 17-19, 1957) New York University and the Radio Corporation of America, 8-13.

[13] Tsao, C. K. (1954). An extension of Massey's distribution of the maximum deviation of two cumulative step functions. Ann. Math. Statist., 25, 587-592.

[14] Wald, A. and Wolfowitz, J. (1944). Statistical tests based on permutations of the observations. Ann. Math. Statist., 15, 358-372.

[15] Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics, 1, 80-83.