Sequential Rank Tests

II. A Modified Two-Sample Procedure

by

Ralph A. Bradley, Sarla D. Merchant
and Frank Wilcoxon

FSU Statistics Report M80
ONR Technical Report No. 16

on

THE DEVELOPMENT OF STATISTICAL METHODS
FOR QUALITY CONTROL AND SURVEILLANCE TESTING

June, 1965
Department of Statistics
Florida State University
Tallahassee, Florida

Research supported by the Army, Navy and Air Force under Office of Naval Research Contract Number NONR-988(08), Task Order NR 042-004 with the Florida State University. Reproduction in whole or in part is permitted for any purpose of the United States Government.
Sequential Rank Tests

II. A Modified Two-Sample Procedure

Ralph A. Bradley, Sara D. Merchant
and Frank Wilcoxon

The Florida State University
Tallahassee, Florida

Summary

This paper describes modifications of the sequential, two-sample, within-group, rank tests developed earlier. A modified, configural rank test is discussed in some detail and is a procedure based on rerankings of observations as new groups of observations are obtained sequentially. A numerical application is given.

Monte Carlo results are presented on the modified, configural rank test and compared with earlier studies. Proofs of termination of the test are reported and outstanding unsolved problems noted.

---

1Research supported by the Army, Navy and Air Force under an Office of Naval Research Contract. Reproduction in whole or in part is permitted for any purpose of the United States Government.

2Present address: Computer Center, Yale University, New Haven, Connecticut.
1. Introduction

In the first paper of this pair, Bradley, Martin and Wilcoxon [3] presented the results of Monte Carlo studies of a sequential, two-sample, grouped, rank-sum test, one of two sequential tests developed earlier [13]. It was noted in [13, p. 74] that a modified ranking system was under consideration and that it should lead to increased efficiencies in sequential, two-sample, rank tests. It is the objective of this paper to present the modified procedures together with some new Monte Carlo results which permit some comparisons with results in [3].

The sequential, two-sample, grouped, rank tests depend on groups of observations taken sequentially with each group consisting of m independent observations from an X-population with cumulative distribution function (cdf) \( F(x) \) and \( n \) independent observations from a Y-population with cdf \( G(y) = F^k(y) \), \( k \geq 1 \). (The form of \( G(y) \) is based on the work of Lehmann [7]). Observations were ranked within each group in joint array and probability ratios for each group and for the set of \( t \) groups at sequential stage \( t \) were obtained for hypotheses, \( H_0: k = 1, F(u) \equiv G(u) \), and \( H_1: k = k_1 > 1, G(u) \equiv F^{k_1}(u) \). The sequential, configural rank test has the probability ratio

\[
\gamma_{(m,n,k_1,1)} = k_1^n \frac{(m+n)!}{(s_1, \gamma-1)!} \prod_{j=1}^{n} \frac{\Gamma(s_j, \gamma-jk_1-1)}{\Gamma(s_j+1, \gamma-jk_1-1)}
\]

(1)

for the \( \gamma \)-th group with \( s_j, \gamma \) being the rank of the \( j \)-th Y-observation in the \( \gamma \)-th group, \( j=1, \ldots, n \), and the complete probability ratio for \( t \) groups of observations is

\[
p_{1t}/p_{0t} = \prod_{\gamma=1}^{t} \gamma_{(m,n,k_1,1)}.
\]

(2)
The corresponding probability ratios for the sequential, rank-sum test are

\[ R_{\gamma}(m,n,k_1,1) = \frac{P(\sum_{j=1}^{n} s_{j,\gamma} = S_{\gamma}|m,n,k_1) / P(\sum_{j=1}^{n} s_{j,\gamma} = S_{\gamma}|m,n,1)} \]

with

\[ P(\sum_{j=1}^{n} s_{j,\gamma} = S_{\gamma}|m,n,k) = \sum_{k=m+n}^{n} \prod_{j=1}^{n} \frac{\Gamma(s_{j,\gamma} + jk - j)\Gamma(s_{j+1,\gamma})}{\Gamma(s_{j+1,\gamma} + jk - j)\Gamma(s_{j,\gamma})} , \]

\[ s_{n+1,\gamma} = m + n + 1, \text{ the sum in (4) being overall sets of } s_{j,\gamma} \text{ with } \sum_{j=1}^{n} s_{j,\gamma} = S_{\gamma}, \text{ and } \]

\[ P_{1t}/P_{0t} = \prod_{\gamma=1}^{t} R_{\gamma}(m,n,k_1,1). \]

The usual Wald bounds apply for the usual sequential divisions;

\[ A = (1-\beta)/\alpha , \quad B = \beta/(1-\alpha), \]

where \(\alpha\) and \(\beta\) are specified probabilities of Type I and Type II errors respectively.

The average sample number (ASN) and operating characteristic functions of the sequential rank tests have been considered through use of Wald formulas in [13] and extensive Monte Carlo investigations have been reported in [3] as already noted. We turn now to modified, sequential, rank tests.

2. Modified Sequential Rank Tests

It appears intuitively that better sequential rank tests might be obtained if complete reranking of the totality of \(X-\) and \(Y-\)observations were effected at each stage of the sequential process. Such a procedure has considerable theoretical
interest and some practical interest, although practical considerations are likely
to dictate within-group ranking in many applications. Merchant [8], working with
Wilcoxon and Bradley, considered this problem.

Suppose that X- and Y-observations are still taken in groups of m and n
and that no group or block effects are present. Then, at the t-th stage of such a
process, mt X-observations and nt Y-observations are ranked in joint array. From
the same assumptions or model as summarized in Section 1, a modified, configural
rank test would be based on the statistic,

\[
p^*_l / p^*_0 = r(mt, nt, k_1, 1)
\]

\[
= k_1^{nt} \frac{((mt)!)\Gamma(s_{1}(t))^{nt} \Gamma(s_{1}(t) + jk_l - j)}{\Gamma(s_{1}(t))^{nt} \Gamma(s_{1}(t) + jk_l - j)},
\]

from (1) wherein \(s_j^{(t)}\) is the rank of the \(j\)th largest Y-observation in the joint
reranking at stage t, \(j=1,\ldots,nt\). Similarly, a modified, rank-sum test may be
based on

\[
p^*_l / p^*_0 = R(mt, nt, k_1, 1)
\]

from (3) and (4), the latter defining \(P(\sum_{j=1}^{nt} s_j^{(t)} = S^{(t)} | mt, nt, k)\).

It was proposed by Merchant [8] that the usual Wald bounds, A and B of
(6), be applied in the modified sequential procedures. That this is appropriate
for the modified, configural rank test has been demonstrated by Savage and Savage
[10] but, for the modified, rank-sum test, no such demonstration is available.

When using a modified, sequential, rank test, we may well choose
\(m = n = 1\). In this way, one should intuitively reduce average sample numbers
slightly but at the cost of more rerankings. The Monte Carlo results reported in
a later section are limited to this situation. Monte Carlo results are not available for the modified, rank-sum test because of the absence of large tables of $R_Y(m,n,k_1,1)$ of (3) for use in (8) and because of the possible inappropriateness of the Wald bounds although an empirical study of use of these bounds would be of interest.

The modified, sequential rank tests seem most useful when rankings are based on quantitative observations. When orderings only are possible, as in experiments involving subjective judgments of individuals, it will usually be necessary to restrict considerations to within-group rankings.

3. An Example

The data on times to death of Control and Experimental animals of [13] are used to illustrate the modified, sequential, configural rank test. While the data were originally presented in four groups, each with $m = n = 5$, we shall suppose for this illustration that the observations were obtained in pairs sequentially as ordered within the groups presented before.

Design parameters are specified as before: $\alpha = .15$, $\beta = .05$, $k_1 = 2.33$, $\ln A = 1.85$, $\ln B = -2.83$. As each new pair of observations is considered, $\ln(p_{1t}^*/p_{0t}^*)$ of (7) is computed and compared with $\ln A$ or $\ln B$. Computations were effected through use of the algorithm described in [13]. Configurations of ordered observations obtained, together with $\ln(p_{1t}^*/p_{0t}^*)$ are shown in Table 1; $X$-observations are on Control animals and $Y$-observations are on Experimental animals.

In the original applications, with within-group ranking, the sequential process terminated with three groups of observations, 15 observations from each population, for the configural rank test. In Table 1, termination occurs with 9 observations from each population and with the decision to accept $H_1: k = 2.33$,
5.

Table 1
Configurations of Ordered Observations and Logarithms of Probability Ratios from Example of Modified Sequential Rank Tests

<table>
<thead>
<tr>
<th>Configurations of Observations</th>
<th>( \ln(p_{1t}^{<em>}/p_{0t}^{</em>}) )</th>
<th>( s(t) )</th>
<th>( \ln(p_{1t}^{<em>}/p_{0t}^{</em>}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>XY</td>
<td>0.3361</td>
<td>2</td>
<td>0.3365</td>
</tr>
<tr>
<td>XYY</td>
<td>0.8150</td>
<td>7</td>
<td>0.8160</td>
</tr>
<tr>
<td>XXXYY</td>
<td>0.9472</td>
<td>14</td>
<td>0.9482</td>
</tr>
<tr>
<td>XXXXYY</td>
<td>0.9751</td>
<td>23</td>
<td>0.9685</td>
</tr>
<tr>
<td>XXXXYXYX</td>
<td>1.5096</td>
<td>36</td>
<td>1.4027</td>
</tr>
<tr>
<td>XXXXYXYXYYX</td>
<td>1.1949</td>
<td>46</td>
<td>0.6329</td>
</tr>
<tr>
<td>XXXXYXYXYYXYYYYX</td>
<td>0.5059</td>
<td>58</td>
<td>0.0534</td>
</tr>
<tr>
<td>XXXXYXYXYYXYYYYX</td>
<td>1.3113</td>
<td>80</td>
<td>0.8044</td>
</tr>
<tr>
<td>XXXXYXYXYYXYYYYXYYX</td>
<td>2.0390(^1)</td>
<td>105</td>
<td>1.6454</td>
</tr>
<tr>
<td>XXXXYXYXYYXYYYYXYYXYYX</td>
<td>1.7355</td>
<td>122</td>
<td>------</td>
</tr>
</tbody>
</table>

\(^1\)Sequential process terminates here.

consistent with the earlier decision. Actually, for the tenth pair of observations, it happened that \( Y < X \) with \( X \) large and, given omission of a decision at stage 9, the decision at stage 10 would have led to continuing experimentation. This example leads to considerable savings in observations for the modified, configural test in comparison with the original configural test; this may be typical as seen from comparisons of average sample numbers in Section 4.
The two final columns in Table 1 show $s(t)$, $t=1,\ldots,10$, the rank sums for the Y-observations, and $\ln(p_{1t}^*/p_{0t}^*)$ of (8) for a possible modified, rank-sum test. On the unverified assumption that the Wald bounds apply, the modified, rank-sum test would not have terminated in the given range of $t$ and it appears that $\ln(p_{1t}^*/p_{0t}^*)$ correlates less well with $\ln(p_{1t}^*/p_{0t}^*)$ as $t$ increases. Values of $\ln(p_{1t}^*/p_{0t}^*)$ were read from the Appendix Table of [13], available for $t=1,\ldots,9$.

4. Monte Carlo Results

Because of the paucity of information on properties of the modified, sequential rank tests, Miss Merchant [8] conducted limited Monte Carlo studies of the modified, sequential, configural rank test. These studies yielded values of the Average Sample Numbers (ASN's) and Power Functions as shown in Tables 2 and 3 respectively. Similar studies have been reported [3] for the sequential, grouped, rank-sum test and pertinent values are included in the tables for comparisons. ASN's are shown in terms of average numbers of observations from each population as only cases with $m=n$ are considered.

Computer programming for Miss Merchant's studies was much as described in [3] with obvious modifications. Basically, the computer was programmed to simulate the sequential experiment so that X- and Y-observations were generated in the computer from populations with cdf's $F(x)$ and $F^k(y)$ respectively for values of $k$ indicated in the tables. The algorithm of [13] was used in the computation of $p_{1t}^*/p_{0t}^*$ of (7).

The Monte Carlo studies were limited to modified, configural rank tests with $\alpha = \beta = .05$ and $m = n = 1$, that is, new observations were obtained in pairs with reranking after each new pair of observations as in the example of Section 3. One would suppose that, if $m = n > 1$ for the modified test, ASN's would increase
Table 2

The Average Sample Numbers of the Modified Configural Rank Test
in Comparison with the Grouped Rank-Sum Tests

<table>
<thead>
<tr>
<th>Design</th>
<th>k</th>
<th>$\mu_y$</th>
<th>$\text{ASN}^{1}$: Modified Config. Test</th>
<th>$\text{ASN}^{1}$: Grouped Rank-Sum Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>m=n=2</td>
</tr>
<tr>
<td>$k_1=1.5$</td>
<td>1</td>
<td>0</td>
<td>54.93$^2$</td>
<td>116.20</td>
</tr>
<tr>
<td>$\alpha=\beta=.05$</td>
<td>3</td>
<td>.846</td>
<td>21.93</td>
<td>31.96</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.163</td>
<td>16.66</td>
<td>22.84</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.352</td>
<td>14.93</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.485</td>
<td>14.29</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.586</td>
<td>13.82</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1.668</td>
<td>13.36</td>
<td>---</td>
</tr>
<tr>
<td>$k_1=2.33$</td>
<td>1</td>
<td>0</td>
<td>19.33</td>
<td>29.16</td>
</tr>
<tr>
<td>$\alpha=\beta=.05$</td>
<td>3</td>
<td>.846</td>
<td>14.78</td>
<td>21.36</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.163</td>
<td>10.12</td>
<td>13.71</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.352</td>
<td>8.84</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.485</td>
<td>8.40</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.586</td>
<td>8.11</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1.668</td>
<td>7.92</td>
<td>---</td>
</tr>
</tbody>
</table>

---

$^1$Average Sample Numbers are average numbers of observations from each population.

$^2$Based on 30 simulated experiments only; all remaining entries based on 500 simulated experiments.
Table 2 -- Continued

<table>
<thead>
<tr>
<th>Design</th>
<th>k</th>
<th>$\mu_y$</th>
<th>ASN: Modified Config. Test</th>
<th>ASN: Grouped Rank-Sum Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>k₁=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8.78</td>
<td>12.46</td>
</tr>
<tr>
<td>$\alpha = \beta = .05$</td>
<td>3</td>
<td>.846</td>
<td>13.09</td>
<td>18.89</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1.163</td>
<td>8.64</td>
<td>11.68</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1.352</td>
<td>6.98</td>
<td>---</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1.485</td>
<td>6.49</td>
<td>---</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1.586</td>
<td>6.34</td>
<td>---</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1.668</td>
<td>5.92</td>
<td>---</td>
</tr>
<tr>
<td>k₁=9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4.53</td>
<td>6.65</td>
</tr>
<tr>
<td>$\alpha = \beta = .05$</td>
<td>3</td>
<td>.846</td>
<td>9.07</td>
<td>12.65</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1.163</td>
<td>7.29</td>
<td>9.54</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1.352</td>
<td>6.05</td>
<td>7.95</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1.485</td>
<td>5.65</td>
<td>7.05</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1.586</td>
<td>5.38</td>
<td>---</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1.668</td>
<td>5.25</td>
<td>---</td>
</tr>
</tbody>
</table>
Table 3

Values of the Power Functions of the Modified Configural Rank Test in Comparison with the Grouped Rank-Sum Tests

<table>
<thead>
<tr>
<th>Design</th>
<th>k</th>
<th>( n_y )</th>
<th>Power: Modified Config. Test</th>
<th>Power: Grouped Rank-Sum Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( m=n=2 )</td>
</tr>
<tr>
<td>( k_1=1.5 )</td>
<td>1</td>
<td>0</td>
<td>.100(^1)</td>
<td>.050</td>
</tr>
<tr>
<td>( \alpha=\beta=.05 )</td>
<td>3</td>
<td>.846</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.163</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.352</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.485</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.586</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1.668</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td>( k_1=2.33 )</td>
<td>1</td>
<td>0</td>
<td>.037</td>
<td>.044</td>
</tr>
<tr>
<td>( \alpha=\beta=.05 )</td>
<td>3</td>
<td>.846</td>
<td>.996</td>
<td>.998</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.163</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.352</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.485</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.586</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1.668</td>
<td>1.000</td>
<td>---</td>
</tr>
</tbody>
</table>

\(^1\)Based on 30 simulated experiments only; all remaining entries based on 500 simulated experiments.
Table 3 -- Continued

<table>
<thead>
<tr>
<th>Design</th>
<th>k</th>
<th>$\mu_y$</th>
<th>Power: Modified Config. Test</th>
<th>Power: Grouped Rank-Sum Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>m=n=2</td>
</tr>
<tr>
<td>$k_1=4$</td>
<td>1</td>
<td>0</td>
<td>.034</td>
<td>.018</td>
</tr>
<tr>
<td>$\alpha=\beta=.05$</td>
<td>3</td>
<td>.846</td>
<td>.898</td>
<td>.900</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.163</td>
<td>.992</td>
<td>.994</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.352</td>
<td>.992</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.485</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.586</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1.668</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td>$k_1=9$</td>
<td>1</td>
<td>0</td>
<td>.034</td>
<td>.048</td>
</tr>
<tr>
<td>$\alpha=\beta=.05$</td>
<td>3</td>
<td>.846</td>
<td>.578</td>
<td>.592</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.163</td>
<td>.884</td>
<td>.900</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.352</td>
<td>.962</td>
<td>.956</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.485</td>
<td>.982</td>
<td>.976</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.586</td>
<td>.990</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1.668</td>
<td>1.000</td>
<td>---</td>
</tr>
</tbody>
</table>
slightly due to the grouping effect. Alternative hypotheses were limited to $k_1 = 1.5, 2.33, 4$ and $9$ corresponding to $p_1 = .6, .7, .8$ and $.9$ where $p_1$ is the probability that a $Y$-observation exceeds an $X$-observation, $p_1 = k_1/(k_1+1)$. Values of $\mu_y$ in the tables are the means of the $Y$-populations given that $F(x)$ is standard normal.

There is confounding present in the comparisons possible from Tables 2 and 3. The modified, configural rank test is compared with the grouped, rank-sum test and not with the grouped, configural rank test. While this may be regrettable, results are not available for the more desirable comparison. However, it has already been indicated in [13] that only small reductions in ASN's result through use of the grouped, configural rank test instead of the grouped, rank-sum test.

Table 2 indicates appreciable savings in observations for the modified, configural rank tests over the grouped, rank-sum tests. Roughly, the reductions in ASN's fall between 20 and 30 per cent. Table 3 shows that the power functions are quite comparable for all sequential tests with values in that table.

While reductions in sample sizes result for the modified, configural rank test, the off-setting effects relate to more computations and increased difficulties in obtaining ranks, the latter perhaps being nearly impossible for some types of experiments. In addition, "block" effects or "trend" effects will disrupt the modified, sequential, rank tests but are easily handled to a large extent in the grouped, sequential, rank tests.

5. Some Theoretical Considerations

The standard theory of sequential analysis gives little information on properties of the modified, sequential, rank tests. The Monte Carlo results of the preceding section provide some assurances as to the efficacy of the modified,
configural rank test but theoretical demonstrations would be most desirable. We report the results of various authors on the theoretical aspects of this work in this section.

Savage and Savage [10] have considered generally the termination of sequential procedures when \( H_0 \) or \( H_1 \) is true, \( k = 1 \) or \( k = k_1 > 1 \) in our problem. Indeed, they indicate application of their theory to our problem when \( m = n = 1 \), that is, when observations are taken in pairs and reranking is effected after each new pair. Thus finite termination and finite expected termination under \( H_0 \) or \( H_1 \) are demonstrated for the modified, configural rank test. Having shown termination and because observation ranks at any stage are sufficient statistics for all observation ranks up to and including a particular stage, it follows from Wald [12, p. 43] that these authors have verified the use of the Wald bounds \( \Lambda \) and \( B \) of (6) for the modified, configural rank test. Termination proofs for the modified, rank-sum test should also be possible but, since the rank-sum at a particular stage is not a sufficient statistic for the whole sequence of rank sums to that stage, it does not follow that the Wald bounds apply.

The extension of the termination proofs to cases with larger \( m \) and \( n \) with reranking after each new group of \( m \) \( X \)-observations and \( n \) \( Y \)-observations follows from the indications in [10] as an easy extension. However, little detail is given there and it seems useful to give a more detailed example.

Theorem 2 of [10] in the notation of that paper is as follows:

"If there exists a sequence \( \{ C_i \} \) of sets with \( C_i \) in the range of \( Y_i \) for which \( E P (\bar{C}_i) \) and \( E P (C_i) \) both converge where \( \bar{C}_i \) is the complement of \( C_i \), then

a. \( \sum \frac{1}{g} N \) is finite,

b. \( \sum \frac{1}{g} g N \) is finite."

It is assumed that \( Y_i \) constitutes the totality of "observations" available at stage \( i \) of a sequential experiment, that \( N \) is the random variable counting the stages \( i \)
to termination, and that \( f \) and \( g \) are an abbreviated notation for \( f_i(Y_i) \) and \( g_i(Y_i) \), the probability functions of \( Y_i \) under \( H_0 \) and \( H_1 \).

Following [10] we let

\[
y_{i\alpha\beta} = 1 \text{ if } x_{\alpha} < y_{\beta} \\
= 0 \text{ if } x_{\alpha} > y_{\beta},
\]

\( \alpha = 1, \ldots, m, \beta = 1, \ldots, n_i \). (Note that the ranks \( s_{(1)}^{(i)}, \ldots, s_{(ni)}^{(i)} \) of the sequence of \( Y \)-observations through stage \( i \) and the corresponding ranks \( t_{(1)}^{(i)}, \ldots, t_{(mi)}^{(i)} \) of the sequence of \( X \)-observations provide information equivalent to that of the \( y_{i\alpha\beta} \)'s.) The sequence \( \{C_i\} \) of sets in the \( Y_i \)-space is defined to be the sequence of regions with

\[
Z_i \leq \frac{(1+3k_1)}{4(1+k_1)}
\]

where

\[
iZ_i = \sum_{j=1}^{i} y_{i(j-1)m+1, (j-1)n+1}.
\]

The model for our sequential rank tests yields \( f_i(Y_i) \) and \( g_i(Y_i) \) such that \( iZ_i \) is equivalent to the number of successes in \( i \) binomial trials with probability of success \( p, p = \frac{1}{2} \) under \( H_0 \) and \( p = k_1/(1+k_1) > \frac{1}{2} \) under \( H_a \). To verify the first condition of the theorem, consider

\[
P_f(C_i) = P_f(Z_i > \frac{1+3k_1}{4(1+k_1)}) = P_f(Z_i - \frac{1}{2} > \frac{k_1-1}{4(1+k_1)}) \\
\leq P_f[\left| Z_i - \frac{1}{2} \right| > \frac{k_1-1}{4(1+k_1)}] \leq K/1^2,
\]

the final inequality resulting from use of the Tchebychev inequality:

\[
P[|X-E(X)| > \varepsilon] \leq E[(X-E(X))^4]/\varepsilon^4. \]

From this bound it follows that \( \sum_{i=1}^{\infty} P_f(C_i) \) converges.
Similarly, $P^\infty(C_i) \leq K' / i^2$ and $\lim_{i \to \infty} P^1(C_i)$ converges. Thus the required termination results follow from the theorem and in addition $P_f(N=\infty), P_g(N=\infty)$ equal zero as indicated in [10].

To complete the demonstration, note that the probability ratio corresponding to the theorem is $P_g(Y_i)/P_f(Y_i)$ which, as noted above, must be equivalent to $P_g(t^{(i)}_{(\alpha)}, s^{(i)}_{(\beta)})/P_f(t^{(i)}_{(\alpha)}, s^{(i)}_{(\beta)})$. But the probabilities for the ranks $t^{(i)}_{(\alpha)}, s^{(i)}_{(\beta)}$ of the ordered X- and Y-observations differ from those of the last probability ratio by the same factor and yield the same probability ratio. Thus $P^*_f / P^*_g$ of (7) is the equivalent probability ratio and the theorem applies to the modified, configural rank test. The appropriateness of the Wald bounds follows for the reasons indicated above.

Other papers and manuscripts bear on our problem. It was noted in [10] that Kraft [6] developed results of a general nature that imply theorems of that paper. Savage and Sethuraman [11] are developing a manuscript showing terminations of the modified, configural rank test for all permissible values of $k$ except possibly for an intermediate value between unity and $k_1$. Their procedures follow, in a more sophisticated way, the approach of Jackson and Bradley [5] and the difficulty with the intermediate value was encountered in that paper but more easily solved.

Hall, Wijsman and Ghosh [4] discuss the modified, configural rank test and note that termination with probability one under $H_0$ or $H_1$ follows from work of Wirjosudirdjo [14] in an unpublished dissertation. Berk has considered the sequential rank tests and has also considered sequential probability ratio tests generally [7]. Parent [9] has included some discussion of termination of the configural rank test but attributes much of the discussion to Savage and Sethuraman while developing a sequential, signed rank test.
Outstanding problems in connection with the modified, sequential, rank tests are:

(i) The development of approximations to ASN-functions.

(ii) The development of approximations to Power functions.

(iii) Verification or refutation of the appropriateness of the Wald bounds for the modified, rank-sum test.

6. Discussion

Initial work on sequential, rank tests seems to have generated considerable research activity in sequential analysis. Extensions of the work of Wald for sequences of dependent observations are becoming available. The contributions of this paper are limited but the paper does serve to set forth the modified, sequential, rank procedures and to present Monte Carlo results indicative of a formal superiority of the modified, configural rank test over the original, configural rank test with within-group ranking. Perhaps a renewed interest in sequential analysis, particularly for nonparametric procedures, will lead to new general theory on ASN- and Power functions.

Acknowledgments

The assistance of National Science Foundation grants to the Florida State University Computing Center is gratefully acknowledged. The courtesy and cooperation of the Computing Center staff made the Monte Carlo studies possible.

Some of the Monte Carlo results of this paper were first presented in [2] at the Fifth International Biometric Conference.
REFERENCES


1. Originating Activity
   Florida State University
   Department of Statistics
   Tallahassee, Florida

2a. Report Security Classification
    Unclassified

2b. Group

3. Report Title
   Sequential Rank Tests II. A Modified Two-Sample Procedure.

4. Descriptive Notes
   Technical Report, June 1965

5. Author(s)
   Bradley, Ralph A., Merchant, Sarla D. and Wilcoxon, Frank

6. Report Date
   June 15, 1965

7a. Total No. of Pages
    16

7b. No. of Refs
    14

8a. Contract or Grant No.
    Nonr 988(08)

b. Project No.

c. Task No.
   NR 042-004

d.

9a. Originator's Report Number(s)
    ONR Technical Report No. 16

9b. Other Report No(s)
    FSU Statistics Report M80

10. Availability/Limitation Notices
    Releasable without limitations on dissemination.

11. Supplementary Notes

12. Sponsoring Military Activity
    Logistics & Mathematical Statistics Branch
    Office of Naval Research
    Washington, D.C. 20360

13. Abstract

   This paper describes modifications of the sequential, two-sample, within-group,
   rank tests developed earlier. A modified, configural rank test is discussed in some
   detail and is a procedure based on rerankings of observations as new groups of
   observations are obtained sequentially. A numerical application is given.

   Monte Carlo results are presented on the modified, configural rank test and
   compared with earlier studies. Proofs of termination of the test are reported and
   outstanding unsolved problems noted.
### Key Words

<table>
<thead>
<tr>
<th>Link A</th>
<th>Link B</th>
<th>Link C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role</td>
<td>Wt</td>
<td>Role</td>
</tr>
</tbody>
</table>

- Sequential Analysis,
- Rank Tests,
- Nonparametric Statistics,
- Statistical Methodology.