

AD-A042 706

WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER
AN ALGORITHM FOR THE DISCRETE FISHER'S PERMUTATION TEST. (U)
APR 77 A P SOMS
MRC-TSR-1745

F/G 12/1

DAA629-75-C-0024
NL

UNCLASSIFIED

| OF |
AD
A042 706



END
DATE
FILMED
8-77
DDC

ADA 042706

MRC TECHNICAL SUMMARY REPORT #1745

AN ALGORITHM FOR THE DISCRETE
FISHER'S PERMUTATION TEST

Andrew P. Soms

Mathematics Research Center
University of Wisconsin-Madison
610 Walnut Street
Madison, Wisconsin 53706

April 1977

(Received February 24, 1977)



Approved for public release
Distribution unlimited

DDC FILE COPY

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

AN ALGORITHM FOR THE DISCRETE FISHER'S PERMUTATION TEST

Technical Summary Report # 1745

ABSTRACT

AMS (MOS) Subject Classification: 62G10

Work Unit Number 4 (Probability, Statistics, and Combinatorics)

[illegible]

Research sponsored by the United States Army under Contract No. DAAG29-75-C-0024, the G. D. Searle & Co., and the University of Wisconsin-Milwaukee.

AN ALGORITHM FOR THE DISCRETE FISHER'S PERMUTATION TEST

Andrew P. Soms^{*}

1. INTRODUCTION

The concept of a permutation or randomization test originated with Fisher [3]. An excellent discussion is given by Conover [2], pp. 357-64, together with extensive references, and the reader is referred to this source for the relevant details. Briefly, let x_i , $1 \leq i \leq k_1$, and y_i , $1 \leq i \leq k_2$, be the observed values from populations 1 and 2, respectively (not necessarily distinct), and let $\bar{x}$ and $\bar{y}$ be the sample means. It is desired to test at level α the hypothesis that populations 1 and 2 are identical against the alternative that population 2 tends to produce smaller values than population 1 (i.e., is stochastically smaller). For Fisher's permutation test, hereafter called the permutation test, the number N of samples of size k_2 that can be drawn from the combined set x_i , $1 \leq i \leq k_1$, y_i , $1 \leq i \leq k_2$, without replacement, with sample mean less than or equal to $\bar{y}$, is counted and the null hypothesis rejected if $N / \binom{k_1 + k_2}{k_2} \leq \alpha$. The permutation Wilcoxon test is carried out in exactly the same way, except that the sum of the average ranks is used in place of $\bar{y}$. These procedures are especially appropriate when there are few distinct values taken on by the sample observations and there are a substantial number of ties in the data.

^{*} Department of Mathematics, University of Wisconsin-Milwaukee, Milwaukee, WI 53201.

Research sponsored by the United States Army under Contract No. DAAG29-75-C-0024, the G. D. Searle & Co., and the University of Wisconsin-Milwaukee.

As an example, in drug screening experiments a control and treatment group are observed for a fixed length of time, at the end of which each subject is assigned an integral numerical score 0, 1, ..., k, with 0 being "best" or normal and k "worst," or completely diseased. It is then desired, using an exact test, to compare the scores of the control group against those of the treatment, to see if the treatment scores are significantly lower.

2. DERIVATION OF THE TESTS

Let each independent control score X_j , $1 \leq j \leq k_1$, have the same density p_i , $P[X_1 = i] = p_i$, $0 \leq i \leq m$, with cumulative distribution function F , and each independent treatment score Y_j , $1 \leq j \leq k_2$, the same density p'_i , $P[Y_1 = i] = p'_i$, $0 \leq i \leq m$, let k_1 and S_1 and k_2 and S_2 be the sample sizes and sample sums for the control and treatment, respectively, and let the total number of 0's, 1's, ..., k's observed be $m_0, \dots, m_k$ (here k is the largest $i \leq m$ such that $m_i > 0$). It is desired to test the null hypothesis $F = G$ against the alternative $F \leq G$, where $F \leq G$ means $F(i) \leq G(i)$, $0 \leq i \leq m$, with at least one strict inequality (i.e., the treatment tends to produce smaller values, or equivalently, Y_1 is stochastically smaller than X_1). Denote by N_i the number of i 's in the treatment, given $m_0, m_1, \dots, m_k$. Then it follows immediately that under the null hypothesis $p_i = p'_i$, $0 \leq i \leq m$,

$$P\left\{N_i = n_i, 0 \leq i \leq k, 0 \leq n_i \leq m_i, \sum_{i=0}^k n_i = k_2\right\} = \frac{k}{\prod_{i=0}^k \binom{m_i}{n_i}} \bigg/ \binom{k_1 + k_2}{k_2}. \quad (1)$$

The tests here described will be conditional, given $m_0, \dots, m_k$. For $k = 1$, (1) is the well known conditional hypergeometric distribution used for the test of the null hypothesis $p_1 = p'_1$ against $p_1 > p'_1$ in two binomial populations with parameters (k_1, p_1) and (k_2, p'_1) (see, e.g., Lehmann [5], pp. 140-3).

The permutation test here described may be regarded as an extension of Fisher's exact test when there are $k+1$ possible outcomes, $0, 1, \dots, k$ ($k \geq 2$), with $i > j$ implying that i is "worse" than j . We used the algorithm to be described below in this particular case and found agreement with the tables in Bennett et al [1]. The distribution in (1) is called the multivariate hypergeometric, and has been discussed by, e.g., Van Eeden [6] and Johnson and Kotz [4], pp. 300-2.

Let the significance level be α . Then for the permutation test we find all those vectors $n = (n_0, \dots, n_k)$ with $\sum_{i=1}^k n_i \leq S_2$, and if the sum of the probabilities of these vectors n is less than or equal to α , we reject the null hypothesis. The procedure for the permutation Wilcoxon test is the same, except that the sum of the average ranks is used. Since this test is invariant under any transformation of the sample values that preserves order, it will always be assumed here that $m_i > 0$, $0 \leq i \leq k$ (if this is not so, the data can be relabelled so that this is true). The average rank corresponding to the value i is $\sum_{j=0}^{i-1} m_j + (m_i+1)/2$ if $i \geq 1$ and $(m_0+1)/2$ if $i = 0$.

For the purpose of deriving a computational procedure for the descriptive level of significance, it is convenient to consider a class of permutation tests of which the two described above are special cases. It is assumed that for each i , $0 \leq i \leq k$, there is a non-negative weight c_i , $c_0 < c_1 < \dots < c_k$, and that the observed value for the treatment is S_2 (i.e., if n_{i0} , $i \leq k$, are the observed treatment frequencies, then $S_2 = \sum_{i=0}^k c_i n_{i0}$). The problem then is to find all vectors $(n_0, \dots, n_k)$ such that $0 \leq n_i \leq m_i$, $\sum_{i=0}^k n_i = k_2$, and $\sum_{i=0}^k c_i n_i \leq S_2$ and sum their probabilities. Since $n_0 = k_2 - \sum_{i=1}^k n_i$, a possible set $n_0, \dots, n_k$ must satisfy

$$\begin{aligned}
c_0(k_2 - \sum_{i=1}^k n_i) + \sum_{i=1}^k c_i n_i &\leq S_2, \\
0 \leq k_2 - \sum_{i=1}^k n_i &\leq m_0, \\
0 \leq n_i \leq m_i, \quad 1 \leq i \leq k.
\end{aligned} \tag{2}$$

The relations (2) are equivalent to

$$\begin{aligned}
n_1 &\leq (S_2 - c_0 k_2 - \sum_{i=2}^k (c_i - c_0) n_i) / (c_1 - c_0), \\
k_2 - m_0 - \sum_{i=2}^k n_i &\leq n_1 \leq k_2 - \sum_{i=2}^k n_i, \\
0 \leq n_1 &\leq m_1, \\
0 \leq n_i &\leq m_i, \quad 2 \leq i \leq k,
\end{aligned} \tag{3}$$

or equivalently,

$$\begin{aligned}
\text{Max}(0, k_2 - m_0 - \sum_{i=2}^k n_i) \leq n_1 \leq \text{Min} \left((S_2 - c_0 k_2 - \sum_{i=2}^k (c_i - c_0) n_i) / (c_1 - c_0), k_2 - \sum_{i=2}^k n_i, m_1 \right), \\
0 \leq n_i \leq m_i, \quad 2 \leq i \leq k.
\end{aligned} \tag{4}$$

In order to continue the process of obtaining limits for $n_2, n_3, \dots, n_k$, in terms of higher subscripted n_i 's, the following fact is needed. For any real numbers $a_i, 1 \leq i \leq n, b_i, 1 \leq i \leq m$,

$$\begin{aligned}
\text{Max}_{1 \leq i \leq n} a_i &\leq \text{Min}_{1 \leq i \leq m} b_i
\end{aligned} \tag{5}$$

if and only if

$$a_i \leq b_j$$

for all i and j . This allows us to obtain intervals for $n_2, n_3, \dots, n_k$. We

give the results for n_2 and n_3 explicitly and also give a general formula for n_j . The interval for n_2 is

$$\begin{aligned} \text{Max}(0, k_2 - m_1 - m_0 - \sum_{i=3}^k n_i) \leq n_2 \leq \text{Min}((S_2 - c_0 k_2 - \sum_{i=3}^k (c_i - c_0) n_i) / (c_2 - c_0) , \quad (6) \\ (S_2 - c_1 k_2 + (c_1 - c_0) m_0 - \sum_{i=3}^k (c_i - c_1) n_i) / (c_2 - c_1) , k_2 - \sum_{i=3}^k n_i , m_2) . \end{aligned}$$

The interval for n_3 is

$$\begin{aligned} \text{Max}(0, k_2 - m_2 - m_1 - m_0 - \sum_{i=4}^k n_i) \leq n_3 \leq \text{Min}((S_2 - c_0 k_2 - \sum_{i=4}^k (c_i - c_0) n_i) / (c_3 - c_0) , \quad (7) \\ (S_2 - c_1 k_2 + (c_1 - c_0) m_0 - \sum_{i=4}^k (c_i - c_1) n_i) / (c_3 - c_1) , \\ (S_2 - c_2 k_2 + (c_2 - c_0) m_0 + (c_2 - c_1) m_1 - \sum_{i=4}^k (c_i - c_2) n_i) / (c_3 - c_2) , k_2 - \sum_{i=4}^k n_i , m_3) . \end{aligned}$$

The general formula for the limits on n_j is given below. The lower limit is given by

$$\text{Max}(0, k_2 - \sum_{i=0}^{j-1} m_i - \sum_{i=j+1}^k n_i) \leq n_j . \quad (8)$$

The upper limit is the minimum of the $j+2$ terms

$$\begin{aligned} k_2 - \sum_{i=j+1}^k n_i , \\ m_j , \quad (9) \\ (S_2 - c_0 k_2 - \sum_{i=j+1}^k (c_i - c_0) n_i) / (c_j - c_0) , \\ (S_2 - c_r k_2 + \sum_{i=0}^{r-1} (c_r - c_i) m_i - \sum_{i=j+1}^k (c_i - c_r) n_i) / (c_j - c_r) , 1 \leq r \leq j-1 . \end{aligned}$$

Note that (4), (6), and (7) are special cases of (8) and (9). For a given value of k , in (8) and (9) $\sum_{i=k+1}^k$ is set to 0 to obtain limits for n_k . The general proof is by induction and follows by observing that in comparing the last $j-1$ terms in (9) to $k_2 - \sum_{i=0}^{j-1} m_i - \sum_{i=j+1}^k n_i$ to obtain the minimum expressions for n_{j+1} , only the term corresponding to $r=j-1$ needs to be retained, since the $j-1$ expressions are non-increasing functions of r .

For the permutation test, $c_i = i$, and for the permutation Wilcoxon test, $c_0 = (m_0+1)/2$ and $c_i = \sum_{j=0}^{i-1} m_j + (m_i+1)/2$ for $i \geq 1$.

3. EXAMPLES

The advantage in obtaining the solution in the above form is that it is ideally suited for a nested do-loop, with n_k being the index for the outermost, n_1 for the innermost loop, the probabilities being summed in the innermost loop. We have written a program for $k=9$ that carries out the permutation and the permutation Wilcoxon test, and the two examples (based on laboratory data) given below were analyzed using this program. Note that if a program has been written for $k=k_0$, it may be used for all $k < k_0$, simply by setting

$m_{k+1}, \dots, m_{k_0}$ equal to 0.

Example 1

In a screening experiment for a drug, $k=4$, $m_0=7$, $m_1=17$, $m_2=19$, $m_3=4$, $m_4=1$, $k_1=25$, $k_2=23$, and $S_2=21$. The actual control results were $n_c = (0, 6, 14, 4, 1)$ and the treatment $n_t = (7, 11, 5, 0, 0)$, where the i^{th} entry is the number of times $i-1$ occurs. Note that if all combinations are examined, there are 28,800 cases if no bounding is done; whereas if the limits (8) and (9) are used, there are only 34 cases to consider. The probability of the set which gives a treatment mean, or equivalently, treatment sum, equal to or smaller than the observed is $.8794 \times 10^{-5}$ and the corresponding probability for the sum of the average ranks is $.9120 \times 10^{-5}$, and thus it is concluded that the

evidence is very strong that there is a causal mechanism depressing the scores of the treatment group.

Example 2

In a screening experiment for a drug, a control and two treatments were used, the control results being $n_c = (0, 0, 10, 13, 2)$, and the treatments $n_{t1} = (1, 5, 11, 6, 0)$ and $n_{t2} = (0, 1, 23, 1, 0)$. Note that the mean score for treatment 1 is 1.96 and for treatment 2, 2.00, the control mean being 2.68. However, for the permutation test the probability associated with treatment 1 is $.1141 \times 10^{-2}$, while for treatment 2 it is $.1067 \times 10^{-4}$, and for the permutation Wilcoxon test the corresponding probabilities are $.1344 \times 10^{-2}$ and $.1067 \times 10^{-4}$. Thus the descriptive levels can go in the opposite direction to the mean scores or sum of average ranks.

4. CONCLUDING REMARKS

The algorithm for the permutation and the permutation Wilcoxon tests here described is most efficient when there are a small number of values taken on by the data. This is also, of course, the situation where the effect of ties is of the greatest concern. For the permutation Wilcoxon test it was shown above that it could be assumed that the data has values $0, 1, \dots, k$ with $m_i > 0$, $0 \leq i \leq k$. For the permutation test, by a change in location and scale (the test is invariant with respect to these transformations), the data can be transformed so that the smallest value is 0 and all the values integral (possibly with gaps). Then the above procedure is applicable if m_i is now defined to be the number of sample values that equal the $i+1^{\text{th}}$ largest data value, with a total of $k+1$ possible, and c_i is the $i+1^{\text{st}}$ largest data value. Thus there is no loss of generality in restricting the sample values to be $0, 1, \dots, k$. We have written a short Fortran computer program, a listing of which is available on request, which carries out the permutation and the

permutation Wilcoxon tests for at most 10 distinct data values. If $\alpha \leq .1$, then in order to have a reasonable running time, the approximate restrictions on the total sample size $k_1 + k_2$ are: for 10 distinct values, ≤ 50 ; for 8, ≤ 80 ; for 6, ≤ 150 ; and for 5, ≤ 250 . Thus if there are 5 or fewer distinct values taken on by the data, then the algorithm here described can be used for all sample sizes encountered in practice. If the data has more than 10 distinct values, then extreme tail probabilities may still be calculated using the above procedure.

The randomized versions of the two tests here discussed are unbiased against the alternatives $F \leq G$. This follows from Lemma 1, p. 73, of [5]. In addition, by using a similar argument to that of [5], pp. 185-8, for the continuous case, it follows that the permutation test is uniformly most powerful in the class of all unbiased tests of $F = G$ against $F \leq G$ for the alternatives $p_i = c(\theta_0)e^{i\theta_0 h(i)}$, $p'_i = c(\theta_1)e^{i\theta_1 h(i)}$, $\theta_1 < \theta_0$, $0 \leq i \leq m$.

In sections 1 and 2 the permutation test and the permutation Wilcoxon test were developed based on population models. They can also be interpreted as randomization tests on the $k_1 + k_2$ experimental units provided that each combination of experimental units has the same probability of being included in the control group. The latter case is, in fact, the most reasonable assumption for the examples discussed.

Many other types of biological experiments also result in data which can be analyzed by the method here proposed when it is desired to compare a treatment to a control and the data has many ties. Some examples are the number of survival days and the number of malformed fetuses in a litter.

References

- [1] Bennett, B. M., Finney, D. J., Hsu, P., and Pearson, E. S., Tables for Testing Significance in a 2x2 Contingency Table, New York: Cambridge University Press, 1963.
- [2] Conover, W. J., Practical Nonparametric Statistics, New York: John Wiley & Sons, Inc., 1971.
- [3] Fisher, R. A., The Design of Experiments, Edinburgh-London: Oliver & Boyd, 1935.
- [4] Johnson, N. L., and Kotz, S., Discrete Distributions, Boston: Houghton-Mifflin, 1969.
- [5] Lehmann, E. L., Testing Statistical Hypotheses, New York: John Wiley & Sons, Inc., 1959.
- [6] Van Eeden, C., "Conditional Limit-Distributions for the Entries in a 2 x k Contingency Table," in G. P. Patil, ed., Classical and Contagious Discrete Distributions, Calcutta, India: Statistical Publishing Society, 1965, 123-6.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER MRC-TSP-1745	2. GOVT ACCESSION NO. 19	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AN ALGORITHM FOR THE DISCRETE FISHER'S PERMUTATION TEST		5. TYPE OF REPORT & PERIOD COVERED Summary Report, no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Andrew P. Soms		8. CONTRACT OR GRANT NUMBER(s) 15 DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 4 (Probability, Statistics, and Combinatorics)
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE April 1977
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 9 (15) 12p.
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) algorithm Fisher's permutation test permutation Wilcoxon test ties		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) An algorithm, especially suitable for a computer, is presented for carrying out Fisher's two-sample permutation test for non-negative integer-valued data. It is shown that the same method can be applied to carry out the permutation Wilcoxon test, using average ranks. Some numerical examples are given and an optimal property of the permutation test is indicated.		