A program for Pitman's permutation test for differences in location

JAMES ARBUCKLE and LEONA S. AIKEN

Department of Psychology, Temple University
Philadelphia, Pennsylvania 19122

Description. One of the earliest and best distribution free tests for differences in location is that of Pitman (1937). Most empirical studies have employed a less desirable test, such as the Mann-Whitney U test, for which tables of critical points exist. The present program applies Pitman's procedure to compute significance levels for an arbitrary set of data. Pitman's test is based upon the principle, enunciated by Fisher (1935), that the conditional probability, given the order statistic, of any permutation of n independent observations is $(n!)^{-1}$. The adaptation of this principle to the two-sample test proceeds as follows. Let $x = (x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_{m+n})$ be $m + n$ observations, where the first m observations come from Sample 1 and the last n observations come from Sample 2; and let $t = [x_{(1)}, \ldots, x_{(m+n)}]$ be the combined order statistic. Let $s = \bar{x}_1 - \bar{x}_2$, where $\bar{x}_1$ and $\bar{x}_2$ are the means in the first and second samples, respectively. The conditional distribution of s given t under the hypothesis that Sample 1 and Sample 2 come from identical populations is obtained by examining every (equally likely) permutation of the observations, computing s for each ordering. The (conditional) probability of any particular value of s, say s_0, is then given by $k(s_0)/(m+n)!$, where $k(s_0)$ is the number of permutations of the observations for which $s = s_0$. Actually, since permutations within Sample 1 and Sample 2 leave s unchanged, it is sufficient to examine the $C(m+n,m) = (m+n)!/(m!n!)$ different partitions of the observations into two groups of sizes m and n.

While this permutation test has many desirable properties (e.g., Kendall & Stuart, 1973), its use has been limited by the fact that the distribution of s is conditional on t, which will change from one occasion to the next, and therefore cannot, in practice, be tabled. One method of avoiding this problem has been to replace the observations by conventional numbers such as rank numbers (the Wilcoxon-Mann-Whitney test), usually sacrificing desirable properties. The present program computes s and assigns a probability level to it by directly examining all partitions of the observations into two groups of sizes m and n.

Input. (1) Title card. (2) Sample sizes. (3) Variable format for data. (4) Observations from first group. (5) Observations from second group.

Output. Output includes the value of s together with one- and two-tailed significance levels.

Limitations. As the program is currently dimensioned, $m + n$ must not exceed 100.

Running time. Running time depends mainly on C(m+n,m) and on min(m,n). By way of illustration, a study with m = 4 and n = 4 required .1 sec while a study with m = 6 and n = 10 required 2.0 sec on the CDC-6400.

Computer and language. This FORTRAN IV program runs in 14 K core on the CDC-6400 operating under SCOPE 3.4.1. No attempt has been made to run the program on other machines, but only standard FORTRAN features have been used. No difficulties in conversion to other machines are anticipated.

Availability. Documentation, source list, and sample output are available without charge from James Arbuckle, Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122.

REFERENCES

Fisher, R. A. *The design of experiments.* London: Oliver and Boyd, 1935.
Kendall, M. G., & Stuart, A. *The advanced theory of statistics* (Vol. 2, 3rd ed.). New York: Hafner, 1973.
Pitman, E. J. G. Significance tests which may be applied to samples from any populations. I. Supplement to the Journal of the Royal Statistical Society, 1937, 4, 119-130.