Six FORTRAN IV programs for conducting simple randomization tests

A. P. WHITE, A. W. STILL, and P. W. HARRIS
Durham University, Durham, England DH1 3LE

Bradley (1968) has shown that, for data that deviate substantially from the assumption of normally distributed error variance required by parametric tests, randomization tests can be more powerful than their parametric counterparts. In addition, randomization tests offer a more finely stepped distribution of the test statistic than do conventional nonparametric tests. This is of particular advantage in cases where group sizes are small, producing very "coarse-grained" distributions of the more commonly used nonparametric test statistics.

Description. This suite of six programs uses several of the combinatorial algorithms from Nijenhuis and Wilf (1975) to carry out two types of test: exhaustive randomization tests and Monte Carlo randomization tests. The exhaustive randomization test calculates values of the test statistic that are obtained by rearranging the given data in all the possible ways consistent with the null hypothesis concerned. From this information, a cumulative probability distribution of the test statistic is derived, against which the observed value can be compared, using either a one-tailed or a two-tailed test. The Monte Carlo randomization test, on the other hand, takes a random sample (with replacement) of the total number of permissible data arrangements in order to derive the distribution of values for the test statistic concerned. The latter method is particularly useful in cases where the total number of permissible arrangements is prohibitively large in terms of the computing time required to perform an exhaustive test. A fuller account of the Monte Carlo procedure is given by Edgington (1969).

The first pair of programs, EMATCH and RMATCH, are the exhaustive and Monte Carlo versions, respectively, of a randomization procedure to test the significance of the difference between the means of two matched samples, $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$. For this test, the null hypothesis is that each matched pair of scores (X_i, Y_i) is drawn from the same population of scores, and hence, the permissible rearrangements of the data, on which the randomization test is based, are those which interchange the positions of the X and Y scores for any number of matched pairs up to a maximum of $n - 1$. Thus, there are 2^{n-1} permissible arrangements of the data, of which the original data are one.

The second pair of programs, EINDEP and RINDEP, perform similar tests for the difference between the means of two independent samples of sizes n_1 and n_2. Here, the null hypothesis is that all the scores are drawn from the same distribution. Thus, the permissible arrangements of the data are those that partition the total set of $(n_1 + n_2)$ scores into two subsets of n_1 and n_2 scores. The number of permissible arrangements is thus $(n_1 + n_2)!/(n_1! \, n_2!)$. As before, one of these is the actual data.

The third pair of programs, ECORR and RCORR, carry out randomization tests to evaluate the significance level of the correlation between two sets of n scores. In this case, the null hypothesis is that the correlation is zero. Hence, the permissible arrangements of the data are those that permute the scores from one set, while holding those from the other stationary. This procedure yields n! permissible arrangements, including the original data. For this test, the sum of the cross products is used as a test statistic, since it necessarily yields exactly the same probability distribution as the correlation coefficient, but is substantially quicker to compute.

Computer and Language. All six programs are written in FORTRAN IV and have been run on an IBM 370/168 computer.

Requirements and Limitations. Four-byte integers are required for some variables.

Availability. Listings and manuals are available at no cost from the first author, Department of Psychology, Durham University, Durham, England DH1 3LE.

REFERENCES

BRADLEY, J. V. *Distribution free statistical tests.* New Jersey: Prentice-Hall, 1968.

EDGINGTON, E. S. Approximate randomization tests. *The Journal of Psychology*, 1969, 72, 143-149.

NIJENHUIS, A., & WILF, H. S. *Combinatorial algorithms.* London: Academic Press, 1975.