

PITMAN
Randomization Tests
Version 5.00S
(c) 1985-1991

"One of many STATOOLS(tm)..."
by

Gerard E. Dallal
54 High Plain Road
Andover, MA 01810

PITMAN performs one- and two-sample exact randomization tests. It also performs Wilcoxon signed-rank tests and Wilcoxon-Mann-Whitney U tests in the presence of an arbitrary number of ties in the data. [The program is available at <http://www.uvm.edu/~dhowell/methods7/Supplements/pitman.exe> .]

NOTICE

Documentation and original code copyright 1985-1991 by Gerard E. Dallal. Neither PITMAN nor its documentation should be modified in any way without permission from the author.

Please acknowledge PITMAN in any manuscript that uses its calculations. A suitable reference is Dallal, GE (1988), "PITMAN: A FORTRAN Program for Exact Randomization Tests," Computers and Biomedical Research, 21, 9-15.

DISCLAIMER

STATOOLS are provided "as is" without warranty of any kind. The entire risk as to the quality, performance, and fitness for intended purpose is with you. You assume responsibility for the selection of the program and for the use of results obtained from that program.

DESCRIPTION

PITMAN is written in FORTRAN-77 around a translation of functions PP1 and PP3 from Streitberg and Rohmel(1984) which calculate the full randomization (permutation) distribution for a set of data. Significance levels for the randomization test (Pitman, 1937) are obtained by applying these functions to the original data, while the levels of the Wilcoxon signed-rank test and the Wilcoxon-Mann-Whitney U test are obtained by applying the functions to the ranked data. Exact levels of the sign test are obtained by evaluating the binomial distribution directly. Exact levels of the median test are obtained by summing the hypergeometric probabilities of all tables with no greater probability than the one observed (Fisher's exact test). Two sample t tests for samples from populations with unequal variances are carried out using Satterthwaite's approximation.

PITMAN analyzes data contained in SYSTAT/MYSTAT files but it cannot make use of character data. If the proposed grouping variable is a character variable, a corresponding numerical variable must be created using MYSTAT's editor.

PITMAN's algorithms can analyze only integer data. Therefore, randomization tests are carried out using only the integer part of the data after multiplication by an appropriate power of 10 to provide as many significant digits as possible given the amount of available storage. (See below.) The translated data are rounded rather than truncated. The test will be equivalent to one based on the original data provided a suitable number of significant digits are used. Size constraints may prove a factor in the analysis of some data sets.

T-tests are carried out using the original data, as are median tests, sign tests, signed rank tests, and Wilcoxon-Mann-Whitney U tests. That is, ranks and signs are determined by the original data rather than the rounded data.

One sample tests require differences. MYSTAT's editor may be used to create a variable containing differences if the data file contains only the original pairs of observations.

PROBLEM SIZE

The maximum number of observations is 100. The shift of the decimal point for randomization tests is determined by the following rules:

Paired data: (sum of absolute values) + 1 .LE. 36,000

Independent samples: If MOBS is the number of observations in the smaller group (here, "observation" is the

difference between the original observation and the smallest observation):

```
((sum of MOBS largest "observations" in the
  combined sample) + largest "observation" + 1)
  * (MOBS + 3) .LE. 36,000
```

and

```
(1 + sum of all "observations") .LE. 36,000
```

If the data are all integers to begin with, PITMAN will not shift decimal points to the right. (Decimal points may be shifted to the left, if necessary.) No shifting to the left takes place for signed rank tests or Wilcoxon-Mann-Whitney U tests; if there is not enough storage, these tests are bypassed and a message is printed. If some ranks are non-integers due to ties, all ranks are doubled.

PAIRED DATA: ZEROS

Zeros are eliminated from the calculations and the sample size is reduced accordingly in all cases except for the t-statistic. Lehmann(1975, pp.144-145) discusses alternate treatments of zeros. There are two additional attractive reasons for discarding zeros: (1) The observed significance level of the randomization test itself is not affected by the presence of an arbitrary number of zeros; (2) Discarding zeros is in keeping with McNemar's test (similar in spirit to the sign test) which disregards ties. The presence of zeros affects significance levels obtained by applying large sample normal theory to the original data, but not as much as one might think. It is instructive to compute a t-statistic for a set of data, add an arbitrary number of zeros, and reanalyze.

For similar reasons, values equal to the median are disregarded when performing a two-sample median test.

ANALYZING A SINGLE SAMPLE

PITMAN can be used to test whether a single sample drawn from a symmetric population has a mean of zero by analyzing the sample as if it were a set of paired data. To test whether the parent population has a mean equal to a specific non-zero value, subtract that value from each observation and analyze the sample as if it were a set of paired data. See the discussion above regarding the treatment of zeros.

TESTS OF DISPERSION

To test whether two populations have different

dispersions, subtract each sample's median from each of its observations and apply PITMAN to the absolute differences.

P-VALUES

There is no universally accepted definition of two-sided P-values for discrete, asymmetric reference distributions. See Gibbons and Pratt (1975) and Yates (1984) for a discussion of the issues. Some possible approaches are:

- (4) Report only one-sided P-values.
- (1) Double the one-sided P-value.
- (3) Sum the probabilities from either tail of the null distribution of the test statistic with likelihood less than that of the observed value.
- (2) Sum the probabilities from either tail of the null distribution of the test statistic corresponding to values that are at least as far from the mean value of the test statistic as the observed value, without regard to their likelihood under the null distribution. This is the approach used by Pitman (1937) and PITMAN.

COMMAND STRUCTURE

This version of PITMAN analyzes SYSTAT/MYSTAT data files. Only the first two letters of a command need be typed.

USE <filename> selects a SYSTAT/MYSTAT system file for analysis.

OUTPUT <filename> directs output to a file.

OUTPUT * directs output to the screen.

OUTPUT @ directs output to the printer.

SUBMIT <filename> submits a batch file of commands.

PAIRED <differences> analyzes a set of differences.

PAIRED <var1> <var2> computes and analyzes the paired differences var1 - var2.

INDEPENDENT <var1> [*] <var2> carries out a two sample test of <var1> using <var2> as a grouping variable.

NAMES displays the variable names in the active file.

FILENAME displays the name of the active file.

HELP displays a help screen.

QUIT quits the program.
PITMAN appends the qualifier .SYS to filenames appearing in
USE commands, the qualifier .DAT to filenames appearing in
OUTPUT commands, and the qualifier .CMD to filenames
appearing in SUBMIT commands.

ALGORITHMS

PITMAN makes use of:

Cooper, B.E. (1968), Algorithm AS 3. The integral of
Student's t-distribution. Appl. Statist.,17,19-20.

and the author's FORTRAN translation of

Pike, M.C. and I.D. Hill (1966). Algorithm 291. Logarithm
of the gamma function. Commun. Ass. Comput.
Mach.,9,684.

REFERENCES

Gibbons, Jean D. and John W. Pratt (1975). P-values:
Interpretation and methodology. The American
Statistician,29,20-25.

Lehmann, E.L. (1975). Nonparametrics. San Francisco:
Holden-Day, Inc.

Pitman, E.J.G. (1937). Significance tests which may be
applied to samples from any populations. Suppl. J.R.
Statist. Soc.,4,119-130.

Streitberg, Bernd and Joachim Rohmel (1984). Exact
nonparametrics in APL. APL Quote Quad, 14(#4), 313-325.

Yates, F. (1984). Tests of significance in 2 X 2 contingency
tables. J. R. Statist. Soc. A,147,426-463.