A Computer Program for Nonparametric Analysis of Complete Repeated Measurements from Multiple Samples

Charles S. Davis
Department of Biostatistics
University of Iowa, Iowa City, IA 52242, U.S.A.

Overview

This program for IBM-compatible personal computers calls Schwertman's (1982) subroutine for nonparametric analysis of complete repeated measurements from multiple samples using the multivariate multisample rank sum test (MMRST) and the multivariate multisample median test (MMMT) of Puri and Sen (1971, Section 5.4). The program compares s groups of subjects or experimental units when measurements are obtained at t time points, or under t conditions, from each subject. In contrast to other methods that require parametric assumptions concerning the distribution of the outcome variable, RMNPM is applicable when the response variable is continuous, but not normally-distributed. The user can specify that the analysis be carried out using the MMRST, the MMMT, or both tests. The program output includes the chi-square statistic with $t(s-1)$ degrees of freedom and p-value for comparing the groups.

Structure of the Input Data File

The input data file must be a standard text (ASCII) file with no hidden characters or word processing format codes. This file should contain one line per subject (independent experimental unit). The data for each subject must include a group identifier and the values of the outcome variable at each of the multiple time points or measurement conditions. Although each record may contain other variables not to be used in the analysis, such as subject identifier, demographic data, etc., the program is restricted to reading in at most 50 variables per subject. Note that the group identifier and the repeated measurements are not required to be in specific fields.

All data items are read in as real numbers using free format input. Thus, fields must be separated by one or more blanks. Data values less than a user-specified missing data code are interpreted as missing values. Experimental units with missing values are excluded from the analysis.

Specification of Analysis Options

The RMNPM options allow the user to determine input/output modes and to specify the number of groups (maximum of 10), the number of variables to be read in (maximum of 50), the index of the group identifier, the codes defining the groups, the number of time points (maximum of 12) to be used in the calculation of the test statistic(s), the indices for the time points to be used, and a missing value indicator. Note, however, that subjects with missing values at any of the t time points are excluded from the analysis. Table 1 further describes the analysis options.

When RMNPM is invoked, the user is asked:

```
ARE INPUT PARAMETERS INCLUDED AT THE BEGINNING OF YOUR DATA FILE?
(1=YES, 2=NO)
```

The simplest method of specifying the desired options is to respond with a '2' and then enter the options interactively in response to program prompts. When options are specified in this manner, the program checks for validity and consistency with the values of previously-entered options. If errors are detected, the user is prompted to reenter the option.

It is sometimes more convenient to include the analysis options at the beginning of the input data file. In this case, options B, C, D, F, H, I, J, and K should be specified on line 1 of the input file, option E should be specified on line 2, and option G on line 3.

It is also possible to include the values of the options in a separate control file. This file must contain one line for each of the 12 options listed in Table 1 and the first line (option A) must contain the value 2. If this file is called `ANALYSIS.CTL`, the command `RMNPM<ANALYSIS.CTL` is then used to invoke the program. If the analysis options are included at the beginning of the data file or in a separate control file, the values of invalid and/or inconsistent options are printed and execution of the program is terminated.

References

Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*. New York: John Wiley and Sons.

Schwertman, N. C. (1982). Algorithm AS 174: Multivariate multisample non-parametric tests. *Applied Statistics*, 31, 80–85.

Table 1
Analysis options

Option	Description
A	Method of specifying input options: 1=included at the beginning of the input file, 2=specified interactively or included in a separate file
B	Number of populations (groups) to be compared (maximum of 10)
C	Total number of variables to be read in (at most 50)
D	Index of the group identifier; permissible values are $1, \ldots, B$
E	The possible values of the group identifier variable D; these define the two groups of subjects
F	Number of time points at which the test statistics are to be calculated; permissible values are $1, \ldots, \min(C - 1, 12)$
G	Indices of variables (time points) at which test statistics are to be calculated; the F values specified must be in ascending order and each must be in the range $1, \ldots, C$ and must not equal D
H	Missing value indicator (data items less than this value are interpreted as missing values)
I	Test statistics to be computed: 1=MMRST only, 2=MMMT only, 3=MMRST and MMMT
J	Code for optional raw data listing: 1=list the data for the time points at which test statistics are to be calculated, 2=do not list the raw data.
K	Name of the input data file (may include a path name)
L	Name of the output data file (may include a path name); enter * if results are to be written to the screen instead of to a file