Behavior Research Methods & Instrumentation
1978, Vol. 10 (3), 417-418

A statistical package for the Hewlett-Packard 2000/Access

MICHAEL D. BIDERMAN
University of Tennessee, Chattanooga, Tennessee 37401

A statistical package that performs analyses commonly used in research and instruction in psychology has been developed for the Hewlett-Packard 2000/Access computer. The package was developed to complement the IDA (for Interactive Data Analysis) package provided with the HP contributed library (Roberts, 1974). Although it can serve as a stand-alone package, it uses the IDA data matrix and common area so that the two packages can be linked, permitting the same data to be analyzed by both.

All analyses performed by the package use a common data matrix. Each analysis is invoked using a command specific to that analysis. After an analysis has been performed, control is returned to a driver program with the data matrix intact. The data matrix is appropriate for small- to medium-sized problems. It can consist of up to 19 columns (variables) by 100 rows (subjects). The number of rows can be increased at the expense of columns.

The package was developed primarily as an instructional tool for undergraduate statistics. As such, it incorporates several features designed to aid the inexperienced user and to protect against loss of data due to entry of inappropriate information. Many commands support a special response, //help, which, if entered in response to a program query, gives a detailed explanation of what is being requested of the user by that query. Another special response, //back, permits the user to back up to the previous query, allowing correction of mistaken entries. A third special response permits the abortion of a command and return to the driver program with the data matrix intact. This is especially useful when only part of the output of a command is desired. Finally, command prompts are tailored to the characteristics of the data matrix at the time the command is executed. For example, a command will not request the column number to be used for an analysis if the data matrix contains only one column of data.

The package includes commands that perform the following analyses:

(1) Descriptive statistics for any column of the data matrix, including measures of central tendency, variability, a simple frequency distribution of values and their z scores, and a simple histogram with automatic scaling of x-axis labels.

(2) Correlation analysis of any pair of columns, including computation of means, standard deviations, Pearson r, test of a hypothesis on r, computation of ranks of scores within each column, Spearman rho, and a scatter plot of z scores of paired values.

(3) Single-group, independent-group, and correlated-group t tests.

(4) One-way independent-groups analysis of variance. If sample sizes are unequal, both least squares and unweighted means analyses are performed. This command allows post hoc comparisons between means using both the Newman-Keuls and Scheffé procedures (e.g., see Winer, 1971, p. 191).

(5) One-way repeated-measures analysis of variance including both the Newman-Keuls and Scheffé post hoc procedures.

(6) Two-way independent-groups ANOVA for equal sample sizes. Either raw data or previously computed sample statistics can be used as input for the above three analyses of variance and the three t tests.

(7) Analysis of variance with one or two repeated-measures factors and one independent-groups factor.

(8) Mann-Whitney U test.

(9) Kruskal-Wallis one-way analysis of variance by ranks.

(10) One-way and two-way chi square.

(11) Fisher's exact probability test for 2 by 2 tables.

(12) One-group multivariate t test.

(13) Two-group multivariate t test with computation of simultaneous confidence intervals on each variable.

(14) One-way multivariate analysis of variance.

(15) Tests of equality and symmetry of covariance matrices (e.g., see Winer, 1971, p. 595).

(16) Principal components analysis (e.g., see Mulaik, 1972, p. 174) with varimax rotation of any number of factors, plotting of any number of pairs of variables in either the unrotated or rotated factor space, and the option of appending factor scores to the data matrix or printing them directly to a data file. This command uses either raw data or correlation coefficients in the data matrix. If correlation coefficients are used, factor score coefficients may be computed but factor scores are not generated.

(17) Nonmetric multidimensional scaling using the method of Johnson (1973). This command includes the options of plotting stimulus coordinates, varimax rotation of coordinates, and printing intermediate and final coordinates to a data file. It uses a lower triangular matrix of similarities or dissimilarities in the data matrix. Due to the amount of time required for the algorithm, it is useful primarily for small demonstration problems.

Many of the commands that test hypotheses report exact significance levels. In addition, cumulative probabilities of the normal, t, chi-square, and F distributions are available from a separate command. Other commands compute the binomial probability distribution, simulate dropping of bags of coins, simulate

random sampling from a single normally distributed population or from two correlated populations, permit data editing, allow "free-form" entry of data (without requiring the user to specify in advance the number of rows to be entered), generate an automatically formatted and paginated display of any subset of the data matrix, and allow programmatic creation and purging of data files.

Language and Computer. The package was written in Hewlett-Packard 2000/Access BASIC (Hewlett-Packard, 1975) and runs on the HP 2000/Access computer at the University of Tennessee at Chattanooga. Modification for use on the HP 3000 appears to be straightforward.

Availability. Cost of a dump of the source code on magnetic tape at 800 bpi may be obtained from the author, Department of Psychology, University of Tennessee, Chattanooga, Tennessee 37401. Copies of a more elaborate description of the capabilities of the package, including sample output, may be obtained at no charge while supplies last.

REFERENCES

HEWLETT-PACKARD COMPANY. *HP 2000/Access BASIC reference manual.* Cupertino, Calif: Author, 1975.

JOHNSON, R. M. Pairwise nonmetric multidimensional scaling. *Psychometrika*, 1973, **38**, 11-18.

MULAIK, S. A. *The foundations of factor analysis.* New York: McGraw-Hill, 1972.

ROBERTS, H. V. *Conversational statistics.* Cupertino, Calif: Hewlett-Packard Company, 1974.

WINER, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.

(Accepted for publication March 21, 1978.)