MULREG: A multiple regression analysis system for microcomputers

DANIEL COULOMBE
University of Ottawa, Ottawa, Ontario, Canada

The package described here is a complete system for data entry, edition, transformation, and analysis using either standard or stepwise (backward elimination procedure) multiple regression with a maximum of more than 50 variables and the number of subjects limited only by the amount of disk storage space. The system is menu driven to allow for maximum ease of use and reliability of operation. The critical matrix inversion routine performs all computations in double precision to reduce the computation errors to a minimum.

The system includes four programs designed respectively to: (1) create, edit, and transform data; (2) run analysis and output to the video display; (3) run analysis and output to the line printer; and (4) produce a scattergram. Computational procedures follow those described by Pedhazur (1982) and Younger (1979).

Input. Data are entered using a routine called CREATEF3/BAS. Each variable is stored as a separate sequential data file. This routine also allows for complete edition of the data (erasing, adding, changing any data piece), transformation of a data file using a user-defined formula that can include another data file (variable), and some other functions related to the management of disk data file (display directory of data file, print data file content, and purge data files).

Analysis. Three options are provided to the user in the context of data analysis. Using the first option, the user can run the analysis and route the output to the video display. The stepwise analysis and some descriptive statistics are not available, but all other features of the system are provided. The main use of this option is to perform an analysis on a limited number of variables, when speed is of some concern.

With the second option (run analysis and output to the line printer), it is possible to select either standard or stepwise regression analysis. The stepwise procedure used is the backward elimination method. If the stepwise option is selected, the standardized residuals and the predicted scores (Y') are saved to disk on each step for later plotting on the line printer.

The third analysis option allows for the production of scattergrams. With this option, it is possible to obtain a scattergram of the relationship between any pair of variables created by the CREATEF3/BAS routine or variables generated by the analysis program itself (residuals and predicted scores for any given step). This option complements the analysis programs and is of particular interest for the visual analysis of residuals.

Output. The output can be directed either to the video display or to the line printer, according to the option selected. It includes descriptive statistics (mean, standard deviation, minimum, maximum, and range for each variable), correlation and variance-covariance matrix, the sums of squares and cross-product matrix, regression statistics (regression coefficients, standardized coefficients, standard error of coefficients, t values, and associated probabilities), the variance-covariance of regression coefficients, the ANOVA table and multiple correlation coefficient, the increment in regression due to each variable, and the residual analysis (residuals, standardized residuals, serial correlation, Durbin-Watson statistic) and plots of residuals against any variable. When the output is directed to the video display, pauses are automatically inserted at the end of each display page. When the output is directed to the line printer, up to 132 columns will be used. The documentation provides informations for those who use 80-column printers.

Limitations. The maximum number of variables that can be used in the analysis is limited to 50, and the maximum number of subjects is limited to about 1,000. With these limitations, the programs use about 38K of memory. DIM statements can be changed to suit particular needs or if more or less memory is available.

Language and Computer. The programs are written in BASIC and have been developed on a TRS-80 Model I and tested with success on a TRS-80 Modell III. Both computers have 48K RAM, and at least two double-density disk drives. Although the system requires two disk drives, minor modifications make it usable in a single-drive environment. The BASIC dialect used is MICROSOFT BASIC compatible and should be tranferable with only slight modifications to the IBM PC and to other microcomputers.

Availability. Program listings and complete documentation are available for $5.00 (to cover copying and mailing costs); a diskette for TRS-80 Model I or Model III is available for $10.00 (cost of diskette, handling, and shipping). Send requests to Daniel Coulombe, School of Psychology, University of Ottawa, 275 Nicholas, Ottawa, Ontario K1N 6N5, Canada.

REFERENCES

PEDHAZUR, E. J. (1982). *Multiple regression in behavioral research.* New York: Holt, Rinehart & Winston.
YOUNGER, M. S. (1979). *A handbook for linear regression.* North Scituate, MA: Duxbury Press.

The author's mailing address: School of Psychology, University of Ottawa, 275 Nicholas, Ottawa, Ontario K1N 6N5, Canada.