

A FORTRAN program to perform analysis of variance and analysis of cross-products

JAMES E. HAYS

The American University, Washington, D.C. 20016

This program performs the analysis of variance or the analysis of cross-products for complete factorial designs. The latter, when combined with analyses of variance on the dependent variable and covariate, can be used to form an analysis of covariance for situations in which there is only one covariate. Analyses of variance for mixed designs can be accomplished by multiple passes of this program.

The program logic follows closely the standard procedures for calculating the terms in an analysis of variance or covariance (Lindquist, 1953). The program can handle any number of observations per cell, as well as those situations in which the cell frequencies are equal, proportional, or consist of only one observation. The program provides very elaborate output that enables the user to tailor the results of the analysis to his particular needs.

In most applications, a minimal number of final calculations must be carried out by hand. This is necessary so as not to limit the applicability of the program to specific problem types. For example, in the analysis of variance, final calculations of F values must be carried out by hand; for the analysis of covariance, the corrected sums of squares and mean squares for the terms involved in the analysis must be calculated; likewise, for mixed designs, some terms in the final summary table must be computed. For most problems, few of these calculations take more than a few minutes on a desk calculator using procedures set forth in standard statistical texts (e.g., Lindquist, 1953).

Input. The data deck consists of the following cards: a control card, a format card, the observation cards, a blank card, and the component cards. The control card indicates to the program the number of observations, the number of cells, the number of independent variables, the number of variance components to be extracted, the type of analysis, and the number of levels for each independent variable. The format card specifies the format which applies to the observation cards; the observation cards contain the dependent variable and

covariate, as well as appropriate identifying subscripts. There may be only one set of observations per card. A blank card follows the last observation card. The component cards specify the exact variance components that are to be extracted during the analysis; for example, specific main effects, two-way interactions, three-way interactions, etc.

Output. The printed output consists of the following: (1) A complete summary of the input control values, the grand mean, and calculated correction term; (2) a summary of the data, indicating the frequency, sum, and sum of squares for each cell; (3) tables of frequencies, means, variances, and effects for each level of each component. Effects are calculated as deviations of the component level mean from the grand mean. Variance estimates may be used directly for homogeneity of variance tests such as the Hartley F max or Cochran tests. In addition, uncorrected sums of squares are also printed for each component. (4) A detailed summary of the analysis, giving the sum of squares, degrees of freedom, and mean square for all components extracted. All output is labeled and self-explanatory.

Computer and Language. The program, written in FORTRAN IV F level, has been successfully run on both an IBM 1130 and an IBM 370/145DOS. The program takes 48,740 bytes of storage in its present form, but it could be adapted to a smaller capacity facility by reducing the total number of cells permitted.

Restrictions. The program has the following limitations: (1) Designs with no more than eight independent variables may be analyzed; (2) no independent variable may have more than 99 levels (this constraint could be easily changed); (3) there may be no more than 500 total cells (this limit may be changed if more core storage is available); and (4) there may be no more than 256 variance components extracted during the analysis.

Availability. Source deck, test data, sample output, and instructions may be obtained free of charge from James E. Hays, Department of Psychology, The American University, Washington, D.C. 20016.

REFERENCE

- Lindquist, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton-Mifflin, 1953.