ANCVB1 and ANCVB2: Unweighted-means, one- and two-way between-subjects analysis of covariance

WILLIAM L. BACHE III

*Department of Child Development and Family Life
Purdue University, Lafayette, Louisiana 47907*

Analysis of covariance programs are routinely available, but nearly all employ a least-squares solution. An unweighted-means solution (Winer, 1971) is indicated if nonorthogonality is introduced via random or accidental differential attrition in the cells of the design. Furthermore, Winer suggests (personal communication) that an unweighted-means solution is usually more appropriate for classification designs even if cell size differences actually do reflect differences in the corresponding population sizes (e.g., psychotics vs. normals). This is because the researcher is usually interested in drawing inferences regarding the *conditions* underlying or attendant to each level of the classification dimension. Differences in these conditions, however, do not usually result from differences in the sizes of the corresponding populations. Therefore, population size, as such, would not ordinarily be considered a salient characteristic of the independent variable. Consequently, the statistical solution should not allow the computation of variances and marginal means to be affected by differences in the cell frequencies. The least-squares solution, which weights cell means by the cell frequencies, allows this to happen. The unweighted-means solution does not. If the design is orthogonal, these two solutions produce identical results. The results differ increasingly, however, as the departure from cell-size equality or proportionality increases.

Input. Program control cards consist of the following: A title card, a parameter card, cell-size cards, dependent-variable name cards, and format cards. The data deck follows if data are punched on cards. Separate decks of control-plus-data cards may be repeated in a single input deck. The parameter card specifies the position of the covariate in a data record (first/last), the number of dependent variables to be analyzed (maximum = 20), the number of cells in each dimension of the design (maximum = 50 for ANCVB1, 10 for ANCVB2), the number of format cards (maximum = 5), the medium on which the data reside (card vs. disk or tape), and an optional flag for rewinding the input file, if other than cards. For ANCVB1, there is also an optional flag which allows the user to exclude zeroes and/or blanks as missing data values.

Output. (1) Summary of the input control parameters. (2) The individual cell frequencies and the harmonic cell frequency. (3) The name as well as the serial number of the dependent variable being analyzed. (4) Cell means for the covariate and the dependent variable, and (for ANCVB1) adjusted cell means for the dependent variable. (5) Various intermediate unadjusted and adjusted variation and covariation terms used in the computation. These would be of help primarily to those following Winer's text. (6) Separate side-by-side analyses of variance and of covariance, with exact probabilities of the chance occurrence of the *F* ratios. (7) With ANCVB1, the error term to be used for hand-calculated multiple comparisons among adjusted cell means (Newman-Keuls, Tukey *a. etc.*). The formula for obtaining this term was adapted from that used in the case of equal cell frequencies (Winer, 1971, p. 772), in that the harmonic cell frequency has been substituted for the

common cell frequency. This substitution was suggested by Winer.

Computer and language. Both programs are written in FORTRAN IV and have been successfully tested on a CDC 6500 using examples from Winer's text. ANCVB1 has been adapted for an IBM 7044, as well.

Availability. User's documentation, program listing, sample output, and a source deck may be obtained free of charge from William L. Bache III, Parent-Child Development Center, 3300 Freret Street, New Orleans, Louisiana 70115.

REFERENCE

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

A correction for Poynor's (EDSTAT) AOVBW program

WILLIAM L. BACHE III

*Parent-Child Development Center
New Orleans, Louisiana 70115*

EDSTAT-V and EDSTAT-P are companion statistical libraries available from Donald Veldman and Hugh Poynor of the University of Texas. One of Poynor's programs (AOVBW) is a highly flexible analysis of variance routine for mixed between-subjects and repeated-measures designs. Not only does it accommodate up to six independent-variable dimensions and unequal cell frequencies, but a large number of dependent variables may be analyzed in a single run, and either zeroes and/or blanks may be automatically excluded as missing-data values. Another notable feature of AOVBW is that it performs a Winer-type unweighted-means solution to the ANOVA problem (Winer, 1971). As Winer points out, this, rather than the least-squares approach, is the solution of choice when equality or proportionality of cell sizes has been lost due to randomly or accidentally caused differential attrition. Even when different cell sizes reflect different sizes of the corresponding populations, as in many classification designs, an unweighted-means solution is more appropriate, because population size as such is not usually a salient characteristic of the independent variable (Winer, personal communication). (When cell sizes are equal or proportional, the two solutions yield identical results.)

Unfortunately, AOVBW does not return an unweighted-means analysis when there are unequal cell frequencies. The problem arises in the computation of the error variances for the within-subjects terms, which are inflated. The corresponding *F* ratios are thus smaller, and the tests of significance are more conservative than they would be if Winer's solution were followed uniformly. (The between-subjects terms are unaffected.) Tracing AOVBW's logic in conjunction with Winer's text, the general nature of the programming error was detected. Not all of the specific errors have been corrected, however, since the decision tree becomes rather complex as the number of independent variables in the design increases. To date, corrections are available which will cause designs to be handled properly as long as they contain no more than two repeated-measures dimensions. The corrections are upward compatible, in that the old conservative solution is returned when this restriction is exceeded.

Input. No changes have been made in the required input.

Output. An appropriate message is printed indicating whether or not Winer's unweighted-means solution is being used.

Restrictions. The corrections are only effective when no more than two repeated-measures dimensions are involved.

William L. Bache III is now at the Parent-Child Development Center, New Orleans, Louisiana 70115.