M. STUART LYNN, Editor

Programming the Tabular Method of Analysis of Variance for Factorial Experiments

S. G. CARMER,
University of Illinois, Urbana, Illinois*

The ease of programming the tabular method of analysis of variance for complete factorial experiments in a FORTRAN language is demonstrated. In this method, the total sum of squares is partitioned into orthogonal single degree of freedom sums of squares; main effect and interaction sums of squares are then obtained by appropriate pooling of the single degree of freedom sums of squares.

Program segments to accomplish the procedure are presented. Modifications to handle hierarchal designs and replicated experiments are mentioned. A FORTRAN II program for an IBM 7094 is described briefly.

Most computer programs for the analysis of variance of factorial experiments employ the conventional desk calculator methods for computing sums of squares described in many statistical methods texts. However, the tabular method of analysis, presented by Bainbridge et al. [1], for any number of factors each of which has any number of levels, offers several appealing advantages as a computer method for obtaining the necessary sums of squares.

Unlike programs where the conventional methods are employed, the length and complexity of a program based on the tabular method is virtually independent of the number of factors and the number of levels of each. A major portion of the computing involves only the calculation of sums rather than the squaring of sums. Finally, all of the arrays need be subscripted in only one or two dimensions, regardless of the number of factors.

In the tabular method, coefficients for single degree of freedom orthogonal comparisons are used for each factor. With the data arranged in a definite pre-sorted order (level of the first factor changing most frequently, . . . , level of the last factor changing least frequently), the orthogonal coefficients are utilized for what amounts to a regression analysis of the data into orthogonal components. The resulting single degree of freedom sums of squares are then pooled in the appropriate manner to obtain main effect and interaction sums of squares for the analysis of variance table.

Since interest lies in the pooled sums of squares the particular orthogonal coefficients chosen for the single degree of freedom comparisons need not be meaningful, and, in fact, may be chosen for ease of computation and programming. A FORTRAN program segment in which suitable coefficients are generated and used to obtain the necessary sums is shown below. For illustration the coefficients generated for a factor having five levels are shown in Table I (see next page).

The variable definitions are:

NOBS	= total number of observations in the experiment
NFACT	= number of factors in the experiment
K = NN(I)	= number of levels of factor I, I = 1, 2, · · · , NFACT
DIVSOR (I, KK)	= divisor for orthogonal comparison (KK−1) for factor I
X	= one-dimensional array which initially contains the data in a definite order and subsequently contains the results of computations involving the orthogonal coefficients
SUM	= one-dimensional array used as temporary storage of results of computations.

The formal program is:

```
      DO 100 I = 1, NFACT
      K = NN(I)
      IGROUP = NOBS / K
      J = 1
      DO 90 KK = 1, K
      L = K - KK + 1
      N = 1
      M = KK - 1
      IDIV = L
      V = 0.0
      IF (L-K) 91, 92, 92
   91 IDIV = IDIV + L*L
      V = L
   92 DIVSOR (I, KK) = IDIV
      DO 80 KKK = 1, IGROUP
      SUM(J) = 0.0
      DO 70 KKKK = 1, L
      SUM(J) = SUM(J) + X(N)
      N = N + 1
   70 CONTINUE
      SUM(J) = SUM(J) - V*X(N)
      N = N + M
      J = J + 1
   80 CONTINUE
   90 CONTINUE
      DO 95 JL = 1, NOBS
   95 X(JL) = SUM(JL)
  100 CONTINUE
```

* Agronomy Statistical Laboratory

Following the last execution of the loop ending with statement 100, the single degree of freedom sums of squares are computed, and the degrees of freedom and sums of squares for each source of variation are pooled. This procedure is shown in the following program segment for situations where the number of factors, NFACT, is less than or equal to three.

The variable definitions are:

NFACT = as before
NN(I) = as before
X = as before
NS(I) = indicators to determine to which sum of squares a particular mean square is to be added.
SS = one-dimensional array for storage of over-all main effect and interaction sums of squares; the number of locations required is ($2^{\text{NFACT}}-1$); for the three factor situation:
- SS(1) = sum of squares for A main effects
- SS(2) = sum of squares for B main effects
- SS(3) = sum of squares for $A \times B$ interaction
- SS(4) = sum of squares for C main effects
- SS(5) = sum of squares for $A \times C$ interaction
- SS(6) = sum of squares for $B \times C$ interaction
- SS(7) = sum of squares for $A \times B \times C$ interaction.

DF = one-dimensional array for storage of degrees of freedom associated with main effect and interaction sums of squares; DF(KK) contains the degrees of freedom associated with SS(KK).

The formal program is:

```
      KKK = 2**NFACT - 1
      DO 150 KK = 1, KKK
      DF(KK) = 0.0
 150  SS(KK) = 0.0
      IC = NN(3)
      IB = NN(2)
      IA = NN(1)
      J = 0
      DO 200 NC = 1, IC
      NS(3) = NC
      DO 200 NB = 1, IB
      NS(2) = NB
      DO 200 NA = 1, IA
      NS(1) = NA
      J = J + 1
      KK = 0
      DIV = 1.0
      DO 190 I = 1, NFACT
      NL = NS(I)
      DIV = DIV*DIVSOR(I, NL)
      IF (NS(I)-1) 190, 190, 191
 191  KK = KK + 2**(I-1)
 190  CONTINUE
      IF (KK) 192, 194, 192
 192  SS(KK) = SS(KK) + X(J)*X(J)/DIV
      DF(KK) = DF(KK) + 1.0
 194  CONTINUE
 200  CONTINUE
```

Although the above program segment is written for up to three factors, it can easily be extended to more factors. For each additional factor, only three additional statements of the type

```
      IC = NN(3)
      DO 200 NC = 1, IC
      NS(3) = NC
```

are required. Additional modifications to handle hierarchal designs could be made rather simply; for example the sum of squares for B within A is equivalent to the sum of squares for B plus the sum of squares for $A \times B$. In replicated experiments, replications can be treated as a factor thus allowing error terms to be obtained from the interactions with replications, as, for example, in randomized complete block, split plot, and repeated measures designs.

Extensive use of the techniques described above has been made in the author's program "Analysis of Variance of Multi-factor (Factorial) Experiments." The program is written in FORTRAN II for an IBM 7094. Replicated experiments with up to five treatment factors and up to 8000 observations can be handled. Up to 10 different sets or kinds of data from a single experiment may be analyzed during one run on the computer. Provisions for numerous transformations of the original data are included in the program. Typical problems involving 30 to 50 treatments and four to six sets of data use approximately 15 to 20 seconds of IBM 7094 time for execution. Output includes a printout of treatment means, analysis of variance table, and (optional) meaningful single degree of freedom comparisons for each set of data; punchout of treatment means is optional. Copies of the source program and a detailed set of instructions are available from the author.

TABLE I. COEFFICIENTS AND CORRESPONDING DIVISORS FOR SINGLE DEGREE OF FREEDOM ORTHOGONAL COMPARISONS FOR A FACTOR WITH $K = 5$ LEVELS AS GENERATED AND UTILIZED IN COMPUTER PROGRAM

Orthogonal comparison	*Values of* *KK*	*M*	*L*	*Coefficients*					*Divisor*
0	1	0	5	+1	+1	+1	+1	+1	$L = 5$
1	2	1	4	+1	+1	+1	+1	−4	$L + L^2 = 20$
2	3	2	3	+1	+1	+1	−3	0	$L + L^2 = 12$
3	4	3	2	+1	+1	−2	0	0	$L + L^2 = 6$
4	5	4	1	+1	−1	0	0	0	$L + L^2 = 2$

RECEIVED JANUARY, 1966; REVISED SEPTEMBER, 1966

REFERENCE

1. BAINBRIDGE, J. R., GRANT, ALISON M., AND RADOK, U. Tabular analysis of factorial experiments and the use of punch cards. *J. Am. Stat. Assoc. 51* (1956), 149–158.