15 Two-Way Analysis of Variance

The preceding chapter discussed ways of evaluating a model in which a nominal variable affects an interval dependent variable. The method, called one-way analysis of variance, consists of computing the variability of the group means weighted by sample size and comparing it to the variability within levels of the independent variable. In this chapter the topic is the study of the simultaneous effects of two independent variables, both measured at the nominal level. It will be shown that two-way analysis of variance is a relatively straightforward extension of one-way analysis of variance. Two-way analysis of variance is sometimes referred to as two-way ANOVA.

Factorial Design

Consider the study by Ball and Bogatz (1970) on the effect of the first year of "Sesame Street" on preschool children. In their study, they divided children into four different viewing groups: (a) nonviewers, (b) occasional viewers, (c) moderate viewers, and (d) heavy viewers. They also classified children as either disadvantaged or advantaged on the basis of neighborhood. There are two independent variables: four levels of viewing and two levels of socioeconomic background. One dependent variable that they studied was the number of letters of the alphabet learned during the six months after "Sesame Street" went on the air. This variable will be called *letters learned*.

In Table 15.1 the various combinations of the Ball and Bogatz evaluation of the television program "Sesame Street" are laid out. The rows in the table are the two levels of socioeconomic status: advantaged and disadvantaged. The columns are the four levels of viewing: none, seldom, moderate, and heavy. Because there are four levels of the viewing variable and two levels of the socioeconomic variable, there are eight possible combinations, also shown in Table 15.1. These combinations are called *cells*. For instance, the

TABLE 15.1 **Factorial Design**

Socioeconomic Status	Viewing: Never	Seldom	Moderate	Heavy	
Advantaged	10	10	10	10	40
Disadvantaged	10	10	10	10	40
	20	20	20	20	80

cell in the upper left-hand corner contains those children who come from advantaged backgrounds and do not watch "Sesame Street." The cell in the bottom right-hand corner contains those children whose parents are disadvantaged and are heavy viewers of "Sesame Street." The creating of all possible combinations is called *factorial design*.

In this chapter, the two nominal independent variables are called factor A and factor B. Factor A has a levels, and factor B has b levels. There are a total of a times b cells in the study. It is usual practice to have an equal number of persons in each of the cells. So, for the "Sesame Street" example in which there are eight cells, if there were 10 children in each cell, there would be a total of 80 children in the study, as is shown in Table 15.1. A table of the n's for the cells is helpful in the computation of two-way analysis of variance.

There are two important reasons for having an equal number of subjects. First, other things being equal, the estimates of the effects of the independent variable are more efficient when the cell sizes are equal. So, to measure more accurately the effect of "Sesame Street," sampling error can be minimized by having equal cell sizes. Second, the computation of the sums of squares becomes much more complicated when the cell sizes are unequal. In fact, there are a number of alternative procedures for estimating the sums of squares. Thus, for reasons of both efficiency and computational ease, equal cell sizes are preferred. All of the discussion in this chapter presumes that cell sizes are equal.

Definitions

The score X_{ijk} refers to score of person i at level j on factor A and at level k on factor B. There are a levels of factor A and b levels of factor B. There are n persons in each cell and a total of ab cells in the design. The total number of scores is abn or N.

To distinguish various summation terms, the following convention will be used in this chapter. The subscript under the summation sign indicates what is summed across. So

$$\sum_{i} X_{ijk}$$

indicates the sum of all the observations in the jk cell,

$$\sum_{i}\sum_{j} X_{ijk}$$

indicates the sum of scores for level k of factor B, and

$$\sum_{i}\sum_{j}\sum_{k} X_{ijk}$$

indicates the sum of all the scores.

Means can be computed for each cell of the design. They are each based on n observations. The cell mean for level j of factor A and level k of factor B is equal to

$$\bar{X}_{.jk} = \frac{\sum_{i} X_{ijk}}{n}$$

The means for factor A are averaged across levels of factor B. Because there are b levels of factor B, there are a total of bn observations that are averaged to compute the a means for factor A. In terms of a formula, the mean for level j of factor A is

$$\bar{X}_{.j.} = \frac{\sum_{i}\sum_{k} X_{ijk}}{bn}$$

The means for factor B are averaged across levels of factor A. Because there are a levels of factor A, there are a total of an observations that are averaged to compute the b means for factor B. In terms of a formula, the mean for level k of factor B is

$$\bar{X}_{..k} = \frac{\sum_{i}\sum_{j} X_{ijk}}{an}$$

The grand mean is denoted as $\bar{X}_{...}$ and it equals the sum of observations divided by the total number of observations; that is,

$$\bar{X}_{...} = \frac{\sum_{i}\sum_{j}\sum_{k} X_{ijk}}{abn}$$

The means for the levels of factor A, the means for levels of factor B, and the grand mean can be expressed in terms of the cell means. The grand mean can be shown to equal

1. the sum of the cell means divided by ab,
2. the sum of the means for factor A divided by a, or
3. the sum of the means for factor B divided by b.

All of these formulas should yield the same value for the grand mean, and so they can be used as a computational check.

The mean for the level j for factor A can be computed by averaging all cell means at level j. There are b such means. The mean for the level k for factor B can be computed by averaging all the cell means at level k. There are a such means.

It is helpful at times to present the cell means, the means for factors A and B, and the grand mean all in one table. Such a table is illustrated in Table 15.2 for the "Sesame Street" example. The numbers in the table are only hypothetical data.

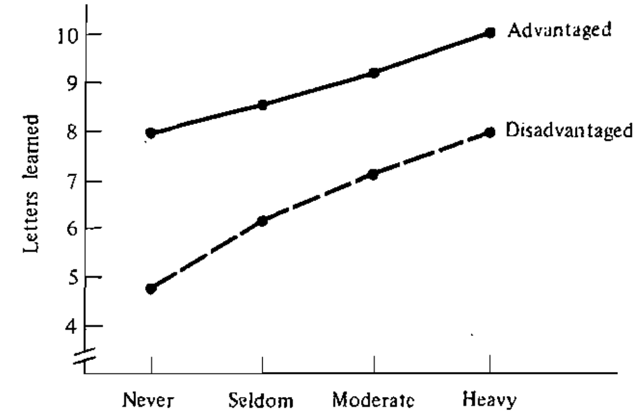
In the last column are the mean for the advantaged group, 8.9 new letters learned, and the mean for the disadvantaged group, 6.7 letters learned. In the bottom row is the set of means for the four viewing groups. They increase from 6.4 to 9.3. In the bottom right-hand corner is the grand mean of 7.8. The entries in the cells are the cell means.

The set of cell means can also be graphed. The factor with the most levels is ordinarily placed on the X axis. In this case that factor would be viewing. The cell means are then plotted on the graph, and one makes certain to place the mean in the appropriate place above the X axis. The points are connected for each level of the second factor (the one not on the X axis). So, as in Figure 15.1, the points for the advantaged groups are connected. To distinguish the two lines one can be solid and the other dashed, as in the figure.

TABLE 15.2 **Hypothetical Table of Means for Two-Way ANOVA**

	Viewing				
Socioeconomic Status	Never	Seldom	Moderate	Heavy	
Advantaged	8.0	8.4	9.2	10.0	8.9
Disadvantaged	4.8	6.2	7.2	8.6	6.7
	6.4	7.3	8.2	9.3	7.8

FIGURE 15.1 **Graph of hypothetical means from the "Sesame Street" evaluation.**



The Concepts of Main Effect and Interaction

One purpose of conducting a two-way analysis of variance is to measure and test the effect of each of the two independent variables. So, with two-way ANOVA, two effects can be tested for the price of one. In two-way analysis of variance, however, there is more than one way to measure the effect of an independent variable. For instance, for the "Sesame Street" example, there is first the program's effect on advantaged children and second its effect on disadvantaged children. The *main effect* of a given independent variable is the effect of that variable averaged across all other levels of the other independent variable. One advantage of having equal numbers of subjects in each cell is that to compute the main effect of one independent variable, one adds the means across cells of the other independent variable and divides by the number of cells to compute the means of a main effect.

To interpret the main effect one examines the means for that factor. Returning to the table of means for the "Sesame Street" example, first the means are examined for the advantaged and disadvantaged subjects. They show that advantaged children learned more letters than disadvantaged children. To determine the main effect for viewing, the four means in the bottom row of Table 15.2 are examined. They show that the more often the children viewed "Sesame Street" the more letters they learned.

Alternatively, one can examine the main effects graphically. Returning to the "Sesame Street" example, it can be seen in Figure 15.1 that the line for advantaged children is always above the line for disadvantaged children. Thus, advantaged children outperform disadvantaged children at all four levels of viewing. Also the lines for both groups increase as the eye moves along the X axis. So, children who view more "Sesame Street" learn more letters.

There are two major purposes in doing a two-way analysis of variance instead of doing separate studies, one for each independent variable. First, it is much more economical to have one study with about the same number of subjects and perform a two-way analysis of variance. One gets two studies for the effort of one. Second, with a two-way analysis of variance one gets information concerning the presence of interaction between the two variables. *Two variables are said to interact if the effect of one variable on the dependent variable varies as a function of the level of the other variable.* Consider the effect of "Sesame Street" on the learning of letters as shown in Table 15.2 and Figure 15.1. If the effect of the program is stronger for lower socioeconomic children than for higher socioeconomic children, it can be said that viewing "Sesame Street" and socioeconomic status *interacted* in causing the learning of letters. This is indicated in both the table and the figure. The effect of "Sesame Street" is greater for disadvantaged than for advantaged children.

The interaction between factor A and factor B is ordinarily referred to as the A by B interaction and it is usually symbolized by $A \times B$.

Predictions of interaction are very common in the social and behavioral sciences. For instance, one question of particular interest is the interaction between diagnostic category and form of therapy. If alcoholics were more helped by group therapy than traditional individual psychotherapy and neurotics were more helped by individual psychotherapy, it would be said that diagnostic category (alcoholic versus neurotic) interacts with mode of therapy (group versus individual).

Another example of interaction might be found when examining the effect of inhaling one milliliter of a toxic drug in the workplace and having a full versus an empty stomach. It might be found that inhaling a toxic drug with a full stomach has relatively little harmful effect, whereas the drug is quite toxic when inhaling it on an empty stomach. In this case, the drug (none versus inhaling one ml) interacts with having eaten (empty versus full stomach) to cause a toxic reaction.

In discussing an interaction, it is said that the effect of factor A on variable X varies depending on the level of factor B. Alternatively, it must also be true that the effect of factor B on variable X varies as a function of the level of factor A. Thus there is a choice in saying which variable's effect changes as a function of which other variable. If the interest is primarily in factor A, then the preference is to state that A's effect changes as a function of B. For instance, for the "Sesame Street" example, instead of saying that the effect of the program was greater for disadvantaged children, it could have been stated that the advantaged children outperformed the disadvantaged children least when both groups were heavy viewers of the program. Also, if A has more levels than B it is probably simpler to say that A's effect changes as function of B.

Interactions can also be represented graphically. The dependent variable is on the Y axis and the independent variable with a larger number of levels is on the X axis. The means of the dependent variable are graphed on the X axis separately for each level of the other independent variable. If the distance between the lines on the graph varies, then an interaction is present. In Figure 15.2 are examples of graphs with interaction and with no interaction. (The diagrams in Figure 15.2 are idealized in that there is no sampling error; actually, graphs ordinarily do not show such clear patterns.)

In both graphs in the figure, six means from a 2 by 3 design are plotted. There are three levels of factor A and two levels of factor B. The X axis is used to distinguish levels of A and two separate lines are drawn for the two levels of B. In the graph labeled I, the gap between the pair of B means increases as one moves along the X axis. It is smallest for A1 and largest for A3. There is then an interaction between the two independent variables. The difference between the B means varies as a function of A. However, in the graph on the right labeled II, the gap remains the same. There is then no interaction between the independent variables. The effect of B is the same for the three levels of A.

To understand better the concept of interaction, examine the graph on the left of Figure 15.3 which is labeled as I. (Again, these are idealized patterns without sampling error.) The graph very clearly shows that A and B interact. At A1, there is no difference between B1 and B2. But as the eyes move to the right on the X axis, the effect of B becomes larger. Although the graph on the left-hand side of Figure 15.3 shows clear interaction, it also shows clear main effects of A and B. For A, the A3 means are on average larger than the A2 and A1 means, and the A2 means are on average larger than the A1 means. For B, the average of the B1 means is larger than the average of the B2 means. So, the pattern of means in the left part of Figure 15.3 shows two main effects and an interaction.

FIGURE 15.2 **Illustration of interaction.**

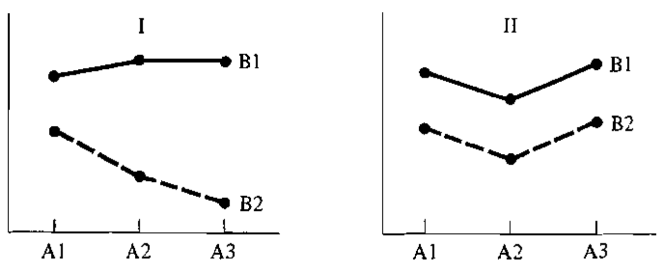
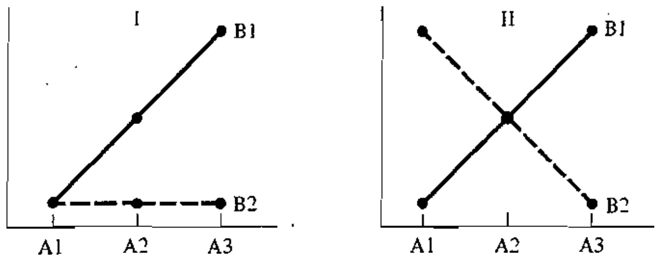


FIGURE 15.3 **Second illustration of interaction.**



In Figure 15.3 the graph on the right shows pure interaction and no main effects. The distinctive feature about the graph is the crossover of the two B lines which must be present when there is interaction and no main effects.

Consider a final example based on actual data. West and Shults (1976) examined how much persons liked male and female names. They determined the commonness of the name by its frequency of occurrence in a college yearbook. They asked 148 persons to state how much they liked common male names such as David and John versus uncommon male names such as Jerome or Julius. They were also asked how much they liked common female names such as Mary and Carol versus uncommon names such as Melinda or Rosemary. There are two factors in this study. They are sex of the name, male or female, and commonness of the name, common or uncommon. The ratings were on a five-point scale, where a five is a favorable rating and a one is an unfavorable rating. The means are presented in Table 15.3.

The results show a main effect of the commonness of the name. Common

TABLE 15.3 **Favorableness of Rating**

Commonness of Name	Sex of Name Male	Female	Average
Common	3.540	3.240	3.390
Uncommon	2.420	2.980	2.700
Average	2.980	3.110	3.045

names are liked more than uncommon names (3.390 versus 2.700). Although male names are liked less than females (2.980 versus 3.110), this difference is too small to be statistically reliable. There is then no main effect for sex of name. But the two factors do interact. Overall, common names are liked better than uncommon names by .69 unit. This effect is strong for male names, a difference of 1.12 units, but the effect is relatively weak for females, a difference of .26 unit. The two factors clearly interact. The effect of commonness depends on gender.

Estimation and Definitional Formulas

The complete model for two-way analysis of variance contains many terms. The dependent variable consists of the following terms:

1. the constant,
2. the main effect for factor A,
3. the main effect for factor B,
4. the interaction between A and B, and
5. the residual variable.

The estimates of these four terms are as follows:

1. The constant: the grand mean or $\bar{X}_{...}$
2. The main effect for factor A at level j: $\bar{X}_{.j.} - \bar{X}_{...}$
3. The main effect for factor B at level k: $\bar{X}_{..k} - \bar{X}_{...}$
4. The interaction between A and B for cell jk:

$$\bar{X}_{.jk} - \bar{X}_{.j.} - \bar{X}_{..k} + \bar{X}_{...}$$

5. The residual score for person i in cell jk: $X_{ijk} - \bar{X}_{.jk}$

The sum of squares for any effect involves the squares of all effects times the sample size that the effect is based on.

It is also necessary to define various totals, as follows:

$$T_{...} = \sum_i \sum_j \sum_k X_{ijk}$$

$$T_{.jk} = \sum_i X_{ijk}$$

$$T_{.j.} = \sum_i \sum_k X_{ijk}$$

$$T_{..k} = \sum_i \sum_j X_{ijk}$$

The F distribution is used to evaluate the plausibility of the restricted models. There are three restricted models in two-way analysis of variance. In each, one of the effects (A, B, or A × B) is omitted.

The Computation of Two-Way Analysis of Variance

A two-way analysis of variance amounts to little more than parts of three separate one-way analyses of variance. So, a sound understanding of one-way analysis of variance is essential for the understanding of two-way analysis of variance.

First, a one-way analysis of variance is computed for factor A ignoring factor B. Second, a one-way analysis of variance is computed for factor B ignoring factor A. Third, a big one-way analysis of variance is computed that treats the cells as levels of a single factor. This last ANOVA can be viewed as an analysis of the AB factor. The computation of two-way analysis of variance consists of the piecing together parts of these three different one-way analyses.

The sums of squares for the main effects of each of the factors ignoring the other are taken from the one-way analyses of variance. The sums of squares interaction is measured by taking the sum of squares from the big one-way analysis and subtracting the sum of squares for both of the main effects.

To compute the sum of squares for A, B, and $A \times B$ the correction term for the mean or C is needed. It equals

$$C = \frac{T_{...}^{\,2}}{abn}$$

In words, it is simply the square of the sum of all scores, divided by the number of all the observations in the study.

To compute the sum of squares for factor A, or SS_A, each A total is squared, these squares are summed across the a groups, this sum is divided by the number of observations that the totals are based on, and C is subtracted. In terms of a formula,

$$SS_A = \frac{\sum\limits_j T_{.j.}^{\,2}}{bn} - C$$

The sum of squares for A using this formula results in the same sum of squares as would be obtained if B were ignored and a one-way analysis of variance sum of squares for A were computed.

To compute the sum of squares for factor B, or SS_B, each B total is squared, these squares are summed across the b groups, this sum is divided by the number of observations that it is based on, and C is subtracted. In terms of a formula,

$$SS_B = \frac{\sum\limits_k T_{..k}^{\,2}}{an} - C$$

To compute the sum of squares for interaction, or $SS_{A\times B}$, the sum of squares AB, or SS_{AB}, is computed. This sum of squares is based on a one-way analysis of variance in which the cell means are treated as group means. Its formula is

$$SS_{AB} = \frac{\sum_j \sum_k T_{.jk}^2}{n} - C$$

The cell totals are divided by n because each cell mean is based on n observations. The formula for the sum of squares for interaction is

$$SS_{A\times B} = SS_{AB} - SS_A - SS_B$$

The sum of squares for subjects within the AB cells is computed indirectly. To compute it, first the sum of squares total, or SS_{TOT}, is computed. Its formula is

$$SS_{TOT} = \sum \sum \sum X_{ijk}^2 - C$$

The sum of squares for subjects within cells equals

$$SS_{S/AB} = SS_{TOT} - SS_{AB}$$

or, alternatively,

$$SS_{S/AB} = SS_{TOT} - SS_A - SS_B - SS_{A\times B}$$

There is a general formula for the sum of squares for a main effect. The formula for the main effect of D is

$$\begin{array}{c} \text{sum of} \\ \text{squares} \\ \text{for effect D} \end{array} = \text{sum}\begin{bmatrix} \text{each of the} \\ \text{D totals} \\ \text{squared} \end{bmatrix} \div \begin{array}{c} n \text{ each} \\ \text{D total} \\ \text{is based on} \end{array} - \begin{array}{c} \text{correction} \\ \text{term for the} \\ \text{grand mean} \end{array}$$

First, each D total is squared and summed across the set of D totals. This sum of totals squared is divided by the sample size of the totals. Finally, C, the correction term for the grand mean, is subtracted.

As was done with one-way analysis of variance, it is useful to diagram the partitioning of the sum of squares. In Figure 15.4 the large circle represents the total sum of squares. The area on inside the two overlapping circles represents the sum of squares AB. The area outside the two overlapping circles represents the $SS_{S/AB}$. It can then be graphically seen that the $SS_{S/AB}$ equals the SS_{TOT} minus the SS_{AB}.

The area in the overlapping circles can be partitioned. The portion of the two smaller circles that does not overlap is the sum of squares for A (on the left) and the sum of squares for B (on the right). Note then that the SS_A is not the entire circle on the left but only the nonoverlapping part.

The part of the two circles that overlaps is the sum of squares interaction. The diagram illustrates how the sum of squares AB equals the sum of squares for the two main effects plus the sum of squares for interaction. Thus the

FIGURE 15.4 **Circle diagram for two-way ANOVA.**

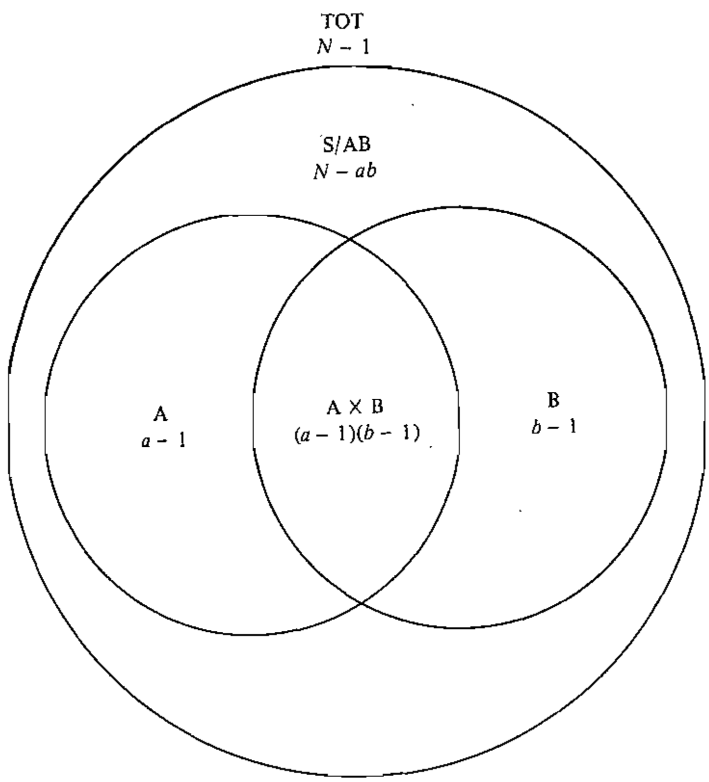


diagram illustrates the partitioning of the sum of squares, as well as their degrees of freedom.

After computing the sum of squares, the next step is to determine the degrees of freedom, as follows:

$$df_A = a - 1$$
$$df_B = b - 1$$

The degrees of freedom for main effect are as they were for one-way ANOVA: They each equal the number of groups less one.

The rule for determining the degrees of freedom for interaction is simple. The *degrees of freedom for interaction* equal the product of the degrees of freedom of its components; that is,

$$df_{A \times B} = (a - 1)(b - 1)$$

The degrees of freedom for subjects within cells equal

$$df_{S/AB} = N - ab$$

where N is the total number of persons in the study. As with the sum of squares, the degrees of freedom partition in the same way:

$$df_{S/AB} = df_{TOT} - df_{AB}$$
$$df_{A \times B} = df_{AB} - df_{A} - df_{B}$$

The df_{AB} equal $ab - 1$, the number of cells less one.

Each mean square is its sum of squares divided by its degrees of freedom. The individual mean squares are

$$MS_A = \frac{SS_A}{df_A}$$

$$MS_B = \frac{SS_B}{df_B}$$

$$MS_{A \times B} = \frac{SS_{A \times B}}{df_{A \times B}}$$

$$MS_{S/AB} = \frac{SS_{S/AB}}{df_{S/AB}}$$

For A, B, and A $\times$ B, the denominator of the F ratio is $MS_{S/AB}$:

$$F(a-1, N-ab) = \frac{MS_A}{MS_{S/AB}}$$

$$F(b-1, N-ab) = \frac{MS_B}{MS_{S/AB}}$$

$$F((a-1)(b-1), N-ab) = \frac{MS_{A \times B}}{MS_{S/AB}}$$

These F tests evaluate whether a restricted model, one that does not include the term in the numerator, is plausible. Note that the df_n differs for these three F tests if there are a different number of levels of A than B. So, in determining the statistical significance of the F tests, different values from Appendix E must be used. A good rule of thumb is that F must be about 4.0 or more to be significant at the .05 level of significance.

As with one-way analysis of variance, the basic results are summarized in an analysis of variance table (see Table 15.4). The column headings for two-way analysis of variance are the same as those for one way analysis of variance. They are source of variation, sum of squares, degrees of freedom, mean square, and F. The sources of variation are factor A, factor B, the A by B interaction, subjects within the AB cells, and total.

TABLE 15.4 **Analysis of Variance Table**

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Factor A	SS_A	df_A	MS_A	$\frac{MS_A}{MS_{S/AB}}$
Factor B	SS_B	df_B	MS_B	$\frac{MS_B}{MS_{S/AB}}$
A × B	$SS_{A\times B}$	$df_{A\times B}$	$MS_{A\times B}$	$\frac{MS_{A\times B}}{MS_{S/AB}}$
Subjects within cells (S/AB)	$SS_{S/AB}$	$df_{S/AB}$	$MS_{S/AB}$	
Total (TOT)	SS_{TOT}	df_{TOT}		

The analysis of variance table neatly summarizes the computation and the results of the model testing.

Assumptions and Power

The assumptions of two-way analysis of variance are identical to those of one-way analysis of variance and the two-sample t test. They are (a) normally distributed residual variable, (b) equal variances in all the cells, and (c) independence of observations. The reader is referred to Chapter 13 for an extensive discussion of these assumptions.

The considerations for the power in the test of the main effects are the same as in one-way analysis of variance, presented in Chapter 14. The power of the test of interaction is ordinarily not as large as that of the main effect. A typical interaction is one in which the effect of one factor becomes weaker across levels of the other factor. Crossover interactions, as in graph II of Figure 15.3, are not common. Thus an interaction measures not some overall effect but the variation of an effect. Therefore main effects are tested with more power than interactions.

Example

Imagine a researcher who wishes to measure the effect of a cigarette smoking on shortness of breath. She creates three groups of smokers: heavy, light, and none. She then divides these three groups by age: 30s, 40s, and 50s. Her raw data are

Smoking Level	*Age*		
	30s	*40s*	*50s*
Heavy	4, 5, 6	5, 6, 10	7, 9, 11
Light	3, 3, 6	2, 4, 6	3, 5, 4
None	2, 2, 5	5, 3, 4	3, 6, 6

She has three persons in each cell of the design.

Some preliminary tables can simplify both the calculations and interpretation. The table of n's for the study is as follows:

		Age			
		30s	40s	50s	Total
Smoking Level	Heavy	3	3	3	9
	Light	3	3	3	9
	None	3	3	3	9
	Total	9	9	9	27

A table of total scores is as follows:

		Age			
		30s	40s	50s	Total
Smoking Level	Heavy	15	21	27	63
	Light	12	12	12	36
	None	9	12	15	36
	Total	36	45	54	135

These tables of n's and totals are useful in computing various sums of squares. Each total squared will be divided by its corresponding n. The means are as follows:

		30s	Age 40s	50s	Average
	Heavy	5.0	7.0	9.0	7.0
Smoking Level	Light	4.0	4.0	4.0	4.0
	None	3.0	4.0	5.0	4.0
	Average	4.0	5.0	6.0	5.0

The correction term for the mean is $135^2/27 = 675$. The sum of squares for smoking groups is

$$\frac{63^2 + 36^2 + 36^2}{9} - 675 = 54$$

The sum of squares for age is

$$\frac{36^2 + 45^2 + 54^2}{9} - 675 = 18$$

The sum of squares for cells is

$$\frac{15^2 + 21^2 + 27^2 + 12^2 + 12^2 + 12^2 + 9^2 + 12^2 + 15^2}{3} - 675 = 74$$

The sum of squares for interaction is

$$74 - 54 - 18 = 12$$

The sum of each squared observation is 813. The sum of squares for the total is then 813 – 675 or 138. The sum of squares for persons within cells is 138 – 74 = 54. The analysis of variance table is as follows:

Source of Variation	*Sum of Squares*	*Degrees of Freedom*	*Mean Square*	*F*
Smoke (S)	54.0	2	27.0	9.0
Age (A)	18.0	2	9.0	3.0
S × A	12.0	4	3.0	1.0
Persons within SA (P/SA)	54.0	18	3.0	

Only the main effect of smoking is statistically significant, and its significance level is the .001 level. The means show that heavy smokers have more difficulty breathing.

Generalization to Higher-Order Analysis of Variance

The generalization to three- and four-way analysis of variance is straightforward. Again, the independent variables are denoted as A, B, and C. There are *a* levels within factor A, *b* levels within factor B, and *c* levels within factor C. There are *abc* cells, each with *n* persons. The total number of subjects in the study, *abcn*, is *N*.

The sums of squares for the main effects are computed exactly as they are computed in one- and two-way analysis of variance. For each level of a main effect, the total is squared and summed across levels, and this sum of squared totals is divided by the number of observations that each total is based on, and the correction term for the grand mean is subtracted.

To determine the interaction between two factors, three "one-way" analyses are done for the AB, AC, and BC means. These sums of squares are denoted as SS_{AB}, SS_{AC}, and SS_{BC}, respectively. The sums of squares for interaction equal

$$SS_{A\times B} = SS_{AB} - SS_A - SS_B$$
$$SS_{A\times C} = SS_{AC} - SS_A - SS_C$$
$$SS_{B\times C} = SS_{BC} - SS_B - SS_C$$

To determine the $SS_{A\times B\times C}$, first the sum of squares for ABC is computed. This is a one-way sum of squares in which the *abc* cells are treated as groups. The sum of squares for the A × B × C interaction is as follows:

$$SS_{A\times B\times C} = SS_{ABC} - SS_{A\times B} - SS_{A\times C} - SS_{B\times C} - SS_A - SS_B - SS_C$$

The sum of squares of subjects within cells equals

$$SS_{S/ABC} = SS_{TOT} - SS_{ABC}$$

where SS_{TOT} equals the sum of each squared observation minus the correction term for the mean.

The degrees of freedom for the two-way interaction are computed in the usual way. They equal the product of the degrees of freedom of the components. They are then

$$df_{A\times B} = (a - 1)(b - 1)$$
$$df_{A\times C} = (a - 1)(c - 1)$$
$$df_{B\times C} = (b - 1)(c - 1)$$

The degrees of freedom for the A × B × C interaction are

$$df_{A \times B \times C} = (a - 1)(b - 1)(c - 1)$$

Like any interaction, its degrees of freedom are the product of the degrees of freedom of its components. The formula for $df_{S/ABC}$ is

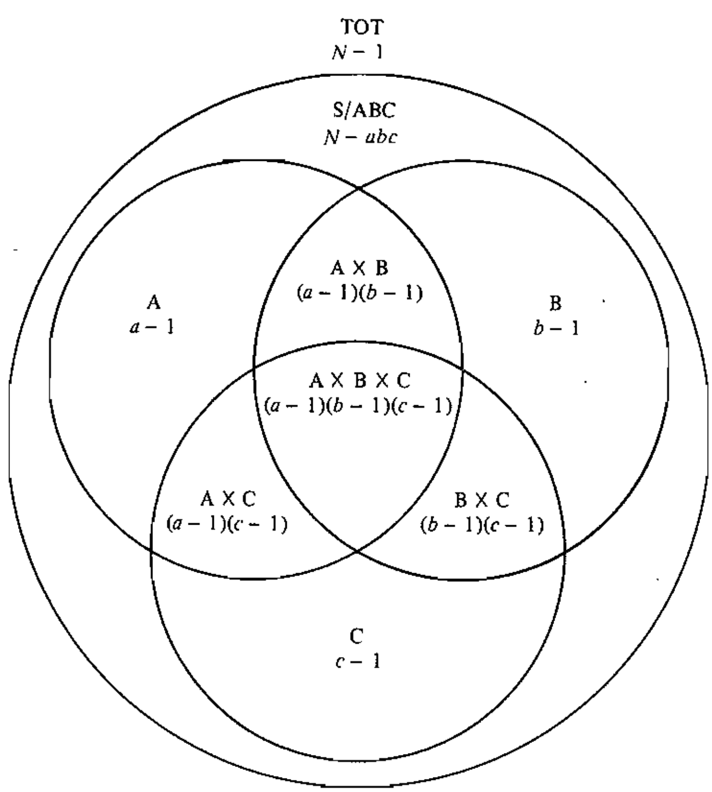
$$df_{S/ABC} = N - abc$$

where N is the total number of observations in the study.

As always, a mean square equals its sum of squares divided by its degrees of freedom. The F test consists of each mean square divided by the mean square subjects within ABC cells.

As was done with two-way analysis of variance, it is useful to diagram the partitioning of the sum of squares. In Figure 15.5 the large circle represents the total sum of squares. The three overlapping circles inside of it represent the sum of squares ABC. The area outside of these three circles, but still within the large circle, represents the $SS_{S/ABC}$. It can then be graphically seen that the $SS_{S/ABC}$ equals the SS_{TOT} minus the SS_{ABC}.

FIGURE 15.5 **Circle diagram for three-way ANOVA.**



The parts of the three circles that do not overlap are the sums of square for A (on the left), for B (on the right), and for C (on the bottom). As with two-way ANOVA, the main effect for an effect is not the entire circle but only the nonoverlapping portion. The portion of the three smaller circles that completely overlaps is the sum of squares interaction of A by B by C. The part of the A and B circles that overlaps excluding C represents the A $\times$ B interaction. In a similar fashion the B $\times$ C and A $\times$ C interactions can be defined. So, the diagram illustrates the partitioning of the sum of squares.

Generalization to a four-way design follows along the same lines. Perhaps the most difficult aspect of four-way designs is that interactions can be very difficult to interpret.

Repeated Measures Design

The estimation procedures for ANOVA have presumed that the groups are independent. In Chapter 13, a design is presented in which observations were matched, paired, or linked in some way. The most common way in which observations are linked is that they come from the same person. For example, consider a small study on the effects of psychotherapy on psychological adjustment. Six subjects were measured before and after psychotherapy on an adjustment scale. Higher scores indicate greater adjustment. The numbers are:

Subject	*Before*	*After*
1	23	32
2	27	25
3	31	40
4	32	31
5	26	38
6	25	29

This is a paired design because one score in each group is linked to the same person.

It is also a two-way design. There are two independent variables. They are psychotherapy, before versus after, and person, 1 through 6. It is possible to compute a mean square subject, a mean square psychotherapy, and a mean square interaction. The two-way analysis of variance table is presented in Table 15.5. The main effect for subject refers to the extent to which subjects differ from one another across both time points. The subject by treatment interaction refers to whether the treatment is more effective for some subjects than others. There is no subjects within cells term that can be used as a denominator for the F test. Instead the person by treatment interaction is used as the denominator for the F test to test for the presence of treatment effects. The value of this F exactly equals the value of t^2 that would be obtained from the paired t test described in Chapter 13. The equivalence of repeated measures F and paired t^2 occurs when the independent variable in the design has two levels.

TABLE 15.4 **Repeated Measures Example**

Source	SS	*df*	MS	*F*
Treatment (T)	80.08	1	80.08	4.80
Person (S)	135.42	5	27.08	
S × T	83.42	5	16.68	
Total	298.92	11		

Wherever there is a series of observations for each person, a two-way analysis of variance can be computed. Such a design is commonly referred to as a *repeated measures design*. This notion of person as variable is a fundamental insight in understanding complicated analyses of variance. Subjects are a very special kind of independent variable, but they are a variable.

Repeated measures designs are very commonly employed in psychological research. In particular, most research in cognitive psychology uses repeated measures designs. Persons in these experiments receive stimuli that are arrayed to represent various levels of a given independent variable. It is not at all uncommon in these studies for a single person to provide data for as many as 25 experimental conditions.

There two major reasons for the popularity of repeated measures designs. First, with a repeated measures design the researcher needs fewer subjects to obtain the same number of observations than is the case with an independent groups design. Second, even if the number of observations is the same for both designs, a repeated measures design is still usually much more powerful than an independent groups design. The reason for this is that subject variation is removed from the residual variance.

Repeated measures designs do have their drawbacks. A complete discussion of these drawbacks can be found in advanced textbooks (Myers, 1979; Winer, 1971). These drawbacks are linked to the fact that measurements are almost always sequentially ordered.

Summary

In two-way analysis of variance, two nominal variables affect a variable measured at the interval level. The creation of all possible combinations of two nominal variables is called *factorial design*. A particular combination in a factorial design is called a *cell*.

The *main effect* of a factor is the average effect of that factor across levels of the other variable. An *interaction* between two factors implies that the effect of one factor changes as a function of the other. An interaction can be assessed by an examination of a graph or table.

Sums of squares for the main effects are computed as in one-way analysis of variance. The interaction sum of squares equals the sum of squares for cells minus the sums of squares for main effects. The sum of squares for subjects within cells is the pooled sum of squares for each cell pooled across cells. This sum of squares is computed by subtracting the sum of squares cells from the sum of squares total. The degrees of freedom for interaction equal the product of the degrees of freedom for the main effects. The mean square for an effect equals its sum of squares divided by its degrees of freedom.

Hypotheses in two-way ANOVA are evaluated by an F test. The denominator of the F ratio is the mean square of subjects within cells. A significant F ratio indicates that the term in the numerator must be included in the complete model.

Repeated measures design involves having each subject be in each level of the independent variable. In a repeated measures design the effect of a factor is tested by using the mean square interaction of subject by factor as the denominator of the F ratio.

Problems

1. Fill in the remainder of the source table from an equal n study with ten subjects in each cell.

Source of Variation	*Sum of Squares*	*Degrees of Freedom*	*Mean Square*	F
A	6.3	3		
B	4.3	2		
A × B	6.0			
Persons within AB (S/AB)				
Total	89.0			

State whether the effects are statistically significant.

2. A pig buyer wants to compare the fat content of bacon and ham in 20 different pigs, 10 from California and 10 from Nebraska. Do a two-way analysis of variance and interpret your results.

	Birthplace	
	Nebraska	*California*
Ham	30, 30, 25, 27, 26	26, 17, 37, 38, 34
Bacon	40, 28, 26, 29, 35	34, 29, 36, 37, 42

3. For the table below do a two-way analysis of variance and interpret your results.

		B 1	2	3	4	5
A	1	7, 9, 6, 5	3, 2, 1, 7	7, 8, 9, 3	1, 3, 8, 7	2, 4, 3, 1
	2	4, 3, 1, 6	4, 1, 3, 5	3, 1, 6, 5	1, 3, 4, 2	7, 6, 5, 9

4. A researcher is interested in studying the overjustification effect. Simply put, this effect states that people do not enjoy activities that they used to do solely for fun after they are paid for engaging in the behavior. A researcher wishes to investigate whether the effect is stronger for younger than for older children. To study the phenomenon the researcher has ten younger children (age four) play with a toy as well as ten older children (age seven). For each of these groups, five of the children were given candy as an incentive to play with the toy and five were not. The experimenter then measured the duration of time spent with the toy at a later period. The results are:

   Older, rewarded: 44, 110, 12, 44, 59

   Older, unrewarded: 79, 120, 112, 68, 39

   Younger, rewarded: 64, 10, 34, 119, 78

   Younger, unrewarded: 73, 10, 102, 49, 99

   Conduct a two-way analysis of variance to see whether the groups differ.

5. Consider the following table of means.

	B1	B2	B3
A1	6.1	5.3	9.0
A2	9.2	8.5	12.1

   Interpret the main effects and interaction.

6. Five subjects are asked to learn material over a period of four days. Their data are

	Day			
Subject	*1*	*2*	*3*	*4*
1	10	15	18	20
2	15	18	18	22
3	10	24	22	25
4	12	19	20	25
5	16	31	40	42

Conduct a two-way ANOVA treating subject and day as factors. Test the effect of day.

7. Complete the following three-way ANOVA table, where there are two subjects in each cell.

SV	*SS*	*df*	*MS*	*F*
A	124	2		
B	25	3		
C	48	1		
A × B	12			
A × C	19			
B × C	14			
A × B × C	12			
S/ABC				
Total	318			

8. Construct a table of means for an experiment in which both independent variables have two levels. Designate the factors as A and B. Make the following three tables.

   a. a table with a main effect for B only
   b. a table with a main effect for A only
   c. a table with an interaction of A with B only

9. An experimenter wants to see whether the deleterious effects of alcohol are increased when one drinks on an empty stomach. He has 20 subjects learn nonsense symbols in one of four conditions: I: no alcohol, empty stomach; II: no alcohol, full stomach; III: alcohol, empty stomach; IV: alcohol, full stomach. He then measured the number correct out of 15 syllables, with the following results.

   I: 14, 12, 15, 14, 13

   II: 13, 15, 14, 12, 14

   III: 8, 7, 5, 9, 12

   IV: 10, 12, 14, 9, 11

   Analyze the data by two-way ANOVA and interpret the results.

10. A researcher wants to see whether intelligence (IQ) is affected by birth order. He finds eight families with three children and measures the IQ of each member. The data are as follows

	Birth Order		
Family	*First*	*Second*	*Third*
1	130	125	120
2	105	90	75
3	140	145	125
4	80	70	80
5	135	120	115
6	90	80	60
7	110	105	90
8	125	100	110

Test whether birth order affects IQ.