7 Factorial Treatment Designs: Random and Mixed Models

The subject of variance component analyses is expanded to more complex designs in this chapter. Variance component models are developed for several variations on the factorial treatment design. Random effects models and models with mixtures of fixed and random effects for factorial arrangements are introduced in this chapter. The concept of the factorial treatment design is extended to include experiments with factors nested within other factors. Designs are discussed for experiments that have a combination of crossed and nested factors. Information for determining replication numbers is included in the discussions. Rules are given for deriving expected mean squares for a variety of balanced factorial experiments.

7.1 Random Effects for Factorial Treatment Designs

Random effects were introduced in Chapter 5 for studies in which the effects of the treatment factor were random samples from a population of treatment effects. The objective was to decompose the total variance into identifiable components. The variability caused by one source or factor can depend on the conditions under which it is evaluated. Thus, some of the total variance is associated with the interaction between two or more factors. The following example illustrates the interaction variance between two factors.

Example 7.1 Evaluating Machine Performance with Variance Components

A manufacturer was developing a new spectrophotometer for use in medical clinical laboratories. The development process was at the pilot stage of

assembly after which machine performance was to be evaluated from assembly line production.

Research Question: A critical component of instrument performance is the consistency of measurements from day to day among machines. In this particular instance, the scientist who developed the instrument wanted to know if the variability of measurements among machines operated over several days was within acceptable standards for clinical applications.

Treatment Design: The scientist set up a factorial treatment design with "machines" and "days" as factors. Four machines were to be tested on four separate days in a 4×4 arrangement.

Experiment Design: Four machines were randomly selected from the pilot assembly production. Eight replicate serum samples were prepared each day from the same stock reagents. Two serum samples were randomly assigned to each of the four machines on each of the four days for a completely randomized design with two replications of each treatment combination. The same technician prepared the serum samples and operated the machines throughout the experiment. The observations on triglyceride levels (mg/dl) in the serum samples are shown in Table 7.1.

Table 7.1 Triglyceride levels (mg/dl) in serum samples run on four machines on each of four days

	Machine			
Day	1	2	3	4
1	142.3, 144.0	148.6, 146.9	142.9, 147.4	133.8, 133.2
2	134.9, 146.3	145.2, 146.3	125.9, 127.6	108.9, 107.5
3	148.6, 156.5	148.6, 153.1	135.5, 138.9	132.1, 149.7
4	152.0, 151.4	149.7, 152.0	142.9, 142.3	141.7, 141.2

Source: Dr. J. Anderson, Beckman Instruments, Inc.

Machines were random factors because they represented a random sample from a potential population of machines to be manufactured, and the Days were a random sample from a population of days on which the machines could be run. The factorial arrangement enabled evaluation of interaction between Machines and Days. The consistency of machine performance would be evidenced by the absence of interaction.

Statistical Model for Variances with Two Treatment Factors

Variability due to the interaction of random factors can play an important role in the inferential process. A random effects model for the two-factor experiment in a completely randomized design is

$$y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk} \qquad (7.1)$$

$$i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, r$$

The random effects a_i, b_j, and $(ab)_{ij}$ are assumed to be independent and normally distributed with means of 0 and variances σ_a^2, σ_b^2, and σ_{ab}^2, respectively. The effects are assumed to be independent of one another. The random errors e_{ijk} are assumed to be independent and normally distributed with mean 0 and variance σ^2.

The observations y_{ijk} in the random effects model have a normal distribution with mean μ and variance

$$\sigma_y^2 = \sigma^2 + \sigma_a^2 + \sigma_b^2 + \sigma_{ab}^2 \qquad (7.2)$$

The components of variance in Equation (7.2) become the focus of any investigation with random effects.

The factorial analysis of variance for the measured concentration of triglycerides in the serum for the spectrophotometer tests described in Example 7.1 is shown in Table 7.2. The analysis of variance computations are those presented in Chapter 6. The expected mean squares are included in the analysis of variance table. A complete set of rules for expected mean square determination applicable to several types of factorial models, including random effects models, is given in Section 7.6.

Table 7.2 Analysis of variance for spectrophotometric readings from four machines on each of four days

Source of Variation	*Degrees of Freedom*	*Mean Square*	*Expected Mean Square*
Day	3	$MSD = 445$	$\sigma^2 + r\sigma_{dm}^2 + rb\sigma_d^2$
Machine	3	$MSM = 549$	$\sigma^2 + r\sigma_{dm}^2 + ra\sigma_m^2$
Interaction	9	$MS(DM) = 87$	$\sigma^2 + r\sigma_{dm}^2$
Error	16	$MSE = 18$	σ^2

Point Estimators for the Variance Components

With equal numbers of observations for each treatment combination the analysis of variance method discussed in Section 5.3 can be used to estimate the components of variance. The analysis of variance is computed for the factorial experiment just as it is for the fixed effects model. The estimates of the components are determined by equating the observed mean squares to the corresponding expected mean squares and solving for the unknown component values. The estimates for the four components of variance in Table 7.2 are

Error: $\hat{\sigma}^2 = MSE = 18.0$

Interaction: $\hat{\sigma}_{dm}^2 = \dfrac{MS(DM) - MSE}{r} = \dfrac{87 - 18}{2} = 34.5$

Machines: $\hat{\sigma}_m^2 = \dfrac{MSM - MS(DM)}{ra} = \dfrac{549 - 87}{2(4)} = 57.8$

Days: $\hat{\sigma}_d^2 = \dfrac{MSD - MS(DM)}{rb} = \dfrac{445 - 87}{2(4)} = 44.8$

The estimate of total variation for a single observation is

$$\hat{\sigma}_y^2 = \hat{\sigma}^2 + \hat{\sigma}_d^2 + \hat{\sigma}_m^2 + \hat{\sigma}_{dm}^2 = 18.0 + 44.8 + 57.8 + 34.5 = 155.1$$

and the estimated standard deviation is $\hat{\sigma}_y = 12.5$.

Tests of Hypotheses About the Variance Components

The significance of the contribution by the components for Machines, Days, and their interaction can be assessed with the F test. The denominator for the F_0 statistic is the mean square with the same expectation as the numerator mean square under the null hypothesis. The test for no interaction, H_0: $\sigma_{dm}^2 = 0$, requires that the mean square for error be used for the denominator since it will have the same expectation as $MS(DM)$ under the null hypothesis. The statistic is

$$F_0 = \frac{MS(DM)}{MSE} = \frac{87}{18} = 4.83$$

and the null hypothesis of no interaction is rejected with $F_0 > F_{.05,9,16} = 2.54$.

The situation is different for tests involving components of variance for main effects. The expected mean squares for main effects are equal to the expected mean square for interaction when the null hypothesis is true. Therefore, the correct F_0 statistic to test H_0: $\sigma_d^2 = 0$ is

$$F_0 = \frac{MSD}{MS(DM)} = \frac{445}{87} = 5.11$$

and the null hypothesis for the Days component is rejected with $F_0 > F_{.05,3,9} = 3.86$. Likewise, the statistic to test H_0: $\sigma_m^2 = 0$ is

$$F_0 = \frac{MSM}{MS(DM)} = \frac{549}{87} = 6.31$$

and the null hypothesis for the Machines component is rejected with $F_0 > F_{.05,3,9} = 3.86$.

An interval estimate for σ^2 can be calculated as shown in Section 5.4 since SSE/σ^2 is a chi-square variable with $ab(r-1)$ degrees of freedom.

Interpretations of the Variance Component Estimates

Each of the components contributes significantly to the variation of a measurement from this particular model of spectrophotometer. The Error component, $\widehat{\sigma}^2 = 18.0$, represents the variation in preparation of the serum samples. The Machines component, $\widehat{\sigma}^2_m = 57.8$, is the variability in machine performance and contributes 37% of the variation. The Days component, $\widehat{\sigma}^2_d = 44.8$, is the variability associated with a new start-up utilizing new reagents for the analysis of samples and other sources of variability that can be identified with day-to-day operational differences. The Interaction component, $\widehat{\sigma}^2_{dm} = 34.5$, contributes 22% of the total variation. The significant interaction implies that the relative performance of the several machines does not vary consistently with the day-to-day changes in the operation. An inconsistency in the calibration of the machines from day to day could be one possible explanation of the interaction.

The factorial design has made it possible to identify several sources of variability in the measurements made by this model of spectrophotometer. The investigator, based on experience, will be able to decide if any of the contributing sources of variability exceeds an acceptable level and correct any deficiencies in the machine or operating conditions if necessary.

Variance Components for Three-Factor Studies

The expected mean squares for a three-factor experiment with random factors are shown in Table 7.3. Some complications arise in the construction of F_0 statistics for tests of hypotheses about the components of variance in random models with more than two factors. The mean square for error can be used to test the hypothesis of no three-factor interaction, and the mean square for three-factor interaction, $MS(\boldsymbol{ABC})$, can be used to test hypotheses for two-factor interaction components. However, upon inspection of the expected mean squares in Table 7.3 it can be seen that there is no legitimate mean square for the denominator of the F_0 statistic to test null hypotheses about the main effect variance components.

Approximate F Tests Required for Some Hypotheses

It is necessary to construct a mean square for the denominator of the F_0 statistic to test the significance of main effect variance components. The construction of approximate F_0 statistics was discussed in Chapter 5 with unequal subsample numbers. For example, to test the hypothesis H_0: $\sigma^2_b = 0$, there are two possible F_0 statistics that approximate the F distribution.

The first approximation is constructed with $MS\boldsymbol{B}$ as the numerator and $M = MS(\boldsymbol{AB}) + MS(\boldsymbol{BC}) - MS(\boldsymbol{ABC})$ as the denominator. The second approximation uses $M1 = MS\boldsymbol{B} + MS(\boldsymbol{ABC})$ as the numerator and

Table 7.3 Expected mean squares for a three-factor experiment with random effects

Source of Variation	*Degrees of Freedom*	*Expected Mean Square*
A	$a-1$	$\sigma^2 + r\sigma^2_{abc} + rc\sigma^2_{ab} + rb\sigma^2_{ac} + rbc\sigma^2_{a}$
B	$b-1$	$\sigma^2 + r\sigma^2_{abc} + rc\sigma^2_{ab} + ra\sigma^2_{bc} + rac\sigma^2_{b}$
C	$c-1$	$\sigma^2 + r\sigma^2_{abc} + rb\sigma^2_{ac} + ra\sigma^2_{bc} + rab\sigma^2_{c}$
AB	$(a-1)(b-1)$	$\sigma^2 + r\sigma^2_{abc} + rc\sigma^2_{ab}$
AC	$(a-1)(c-1)$	$\sigma^2 + r\sigma^2_{abc} + rb\sigma^2_{ac}$
BC	$(b-1)(c-1)$	$\sigma^2 + r\sigma^2_{abc} + ra\sigma^2_{bc}$
ABC	$(a-1)(b-1)(c-1)$	$\sigma^2 + r\sigma^2_{abc}$
Error	$abc(r-1)$	σ^2

$M2 = MS(\boldsymbol{AB}) + MS(\boldsymbol{BC})$ as the denominator. The first ratio may be easier to use because approximate degrees of freedom need be computed for only one of the mean squares by the Satterthwaite procedure (Equation 5.27). However, it is possible to construct a negative mean square when some of the mean squares have negative signs in the function. Gaylor and Hopper (1969) discuss some of the problems associated with approximating the F distribution with linear combinations of mean squares. The recommended F_0 statistic would be the second with synthesized mean squares $M1$ and $M2$.

7.2 Mixed Models

Many experiments are designed to study the effects of one factor on the population mean and the effects of another on the population variance. These experiments have a mixture of fixed factors and random factors. Models for factorial arrangements that include random factors and fixed factors are called **mixed models**, because they contain a mixture of random and fixed effects.

The model and analysis for mixed effects consist of two parts because there are two types of inferences. The inferences for the random effects factor apply to the variation in a population of effects, whereas the inferences for the fixed effects factor are restricted to the specific levels used for the experiment. The experiment described in the following example included a mixture of random and fixed factors.

Example 7.2 Evaluation of Two Chemistry Methods on Four Days

New methods of chemistry frequently are developed to assay for compounds in a clinical laboratory setting. Given the choice among two or more chemistry methods, the clinical chemist must evaluate the relative performance of the methods.

Research Question: In this example, a chemist wanted to know whether the two chemistry methods consistently provided equivalent results in an assay for triglycerides in human serum. A methods comparison test was constructed by the clinical chemist to evaluate the difference in the performance of the two chemistry methods for the assay.

Treatment Design: A factorial treatment design was used with the factors "chemistry methods" and "days." The two methods each were to be tested on four days, for a 2×4 arrangement.

Experiment Design: Each day two replicate samples of serum were prepared with each of the two chemistry methods. The samples were analyzed in random order on the same spectrophotometer for a completely randomized experiment design with two replications. The same technician prepared the serum samples and operated the spectrophotometer for each of the tests. The observations on triglyceride levels (mg/dl) in the serum samples are shown in Table 7.4.

Table 7.4 Triglyceride levels (mg/dl) in serum samples from a factorial arrangement for two chemistry methods, fixed factor, and four days, random factor

	Day			
Method	1	2	3	4
1	142.3, 144.0	134.9, 146.3	148.6, 156.5	152.0, 151.4
2	142.9, 147.4	125.9, 127.6	135.5, 138.9	142.9, 142.3

Source: Dr. J. Anderson, Beckman Instruments, Inc.

Chemistry Method is a fixed factor because the two methods are reproducible in a repeated experiment. The inferences are restricted to a comparison between the two methods used in the experiment. Days, on the other hand, represent random effects because the four days are considered a random sample of days on which the two chemistry methods could be tested.

The advantage of the factorial arrangement in this example is the ability to evaluate the consistency of the two methods under repeated runs. In the absence of interaction the two methods would produce the same relative results from day to day. The presence of interaction would indicate one or both of the methods were inconsistent in repeated applications.

Statistical Model for One Fixed and One Random Factor

The linear model for a two-factor experiment with a fixed effect, factor A, and a random effect, factor B, is

$$y_{ijk} = \mu + \alpha_i + b_j + (ab)_{ij} + e_{ijk} \qquad (7.3)$$

$$i = 1,2,\ldots,a \quad j = 1,2,\ldots,b \quad k = 1,2,\ldots,r$$

where μ is the mean, α_i is the fixed effect for factor A, b_j is the random effect for factor B, $(ab)_{ij}$ is the interaction effect, and e_{ijk} is random experimental error. The random effects, b_j and e_{ijk}, are assumed independent and normally distributed with means 0 and variances σ_b^2 and σ^2, respectively.

The interaction effects $(ab)_{ij}$ are assumed to be random effects, independent and normally distributed with mean 0 and variance σ_{ab}^2. The interaction effects are assumed random when one of the factors involved is a random effect.

Analysis of Mixed-Factor Experiments

The expected mean squares for the mixed-model analysis of variance are different from those for either the completely fixed or completely random effects models. The expected mean squares and the analysis for the factorial experiment with one random and one fixed factor are illustrated for the chemistry method experiment in Example 7.2. The analysis of variance for the 16 observations is shown in Table 7.5. The expected mean squares are included in the analysis of variance table. Note the use of θ_m^2 for the variance of the fixed effect as defined in Table 6.5.

Table 7.5 Analysis of variance for a factorial experiment with one fixed effects factor, Method, and one random effects factor, Day

Source of Variation	*Degrees of Freedom*	*Mean Square*	*Expected Mean Square*
Method	1	$MSM = 329$	$\sigma^2 + r\sigma_{md}^2 + rb\theta_m^2$
Day	3	$MSD = 144$	$\sigma^2 + r\sigma_{md}^2 + ra\sigma_d^2$
Interaction	3	$MS(MD) = 62$	$\sigma^2 + r\sigma_{md}^2$
Error	8	$MSE = 14$	σ^2

Source: Dr. J. Anderson, Beckman Instruments, Inc.

Tests of Hypotheses About Variance Components and Means

The null hypothesis of no interaction, H_0: $\sigma_{md}^2 = 0$, is tested with

$$F_0 = \frac{MS(MD)}{MSE} = \frac{62}{14} = 4.43$$

and the null hypothesis is rejected since $F_0 > F_{.05,3,8} = 4.07$. The presence of interaction with days suggests a possibility of differences among the chemistry methods that vary with days. Comparisons can be made between chemistry methods within each day with contrasts among cell means.

A test of the null hypothesis of no difference between the marginal means of the fixed effect factor would only be appropriate in the absence of interaction. For illustrative purposes only, the F_0 statistic for Chemistry Methods is

$$F_0 = \frac{MSM}{MS(MD)} = \frac{329}{62} = 5.31$$

and the null hypothesis would not be rejected since $F_0 < F_{.05,1,3} = 10.13$. The observed means for the two chemistry methods were $\bar{y}_{1..} = 147$ and $\bar{y}_{2..} = 138$ mg/dl.

The F_0 statistic to test the significance of the variance component for Days, σ_d^2, is

$$F_0 = \frac{MSD}{MS(MD)} = \frac{144}{62} = 2.32$$

and the null hypothesis H_0: $\sigma_d^2 = 0$ would not be rejected since $F_0 < F_{.05,3,3} = 9.28$.

Standard Errors for the Fixed-Factor Means

The standard error for the difference between the two chemistry methods on a given day,

$$\sigma_{(\bar{y}_{ij.} - \bar{y}_{kj.})} = \sqrt{\frac{2(\sigma^2 + r\sigma_{md}^2)}{r}} \qquad (7.4)$$

is estimated by

$$s_{(\bar{y}_{ij.} - \bar{y}_{kj.})} = \sqrt{\frac{2MS(MD)}{r}} = \sqrt{\frac{2(62)}{2}} = 7.9 \qquad (7.5)$$

As a general rule the mean square used for the standard error of a difference between the means of a fixed effect factor is the denominator mean square of F_0 used to test the null hypothesis about the fixed effect. Thus, the standard error of the difference between the marginal means of the two chemistry methods is estimated by

$$s_{(\bar{y}_{i..} - \bar{y}_{k..})} = \sqrt{\frac{2MS(MD)}{rb}} = \sqrt{\frac{2(62)}{2(4)}} = 3.9 \qquad (7.6)$$

Three-Factor Experiments with Random and Fixed Factors

The problem of constructing F_0 statistics for tests of hypotheses was discussed in Section 7.1 for the three-factor experiment with three random factors. Similar difficulties occur with the mixed effects model for experiments with two or more random factors. The expected mean squares are given in Table 7.6 for three-factor experiments that have either one or two random effect factors with the remaining factor(s) fixed.

Table 7.6 Expected mean squares for three-factor experiments with (1) one fixed effect factor and two random effect factors and (2) two fixed effect factors and one random effect factor

	Expected Mean Square	
Source of Variation	*A Fixed, B and C Random*	*A and B Fixed, C Random*
A	$\sigma^2 + r\sigma_{abc}^2 + rc\sigma_{ab}^2 + rb\sigma_{ac}^2 + rbc\theta_a^2$	$\sigma^2 + r\sigma_{abc}^2 + rb\sigma_{ac}^2 + rbc\theta_a^2$
B	$\sigma^2 + r\sigma_{abc}^2 + rc\sigma_{ab}^2 + ra\sigma_{bc}^2 + rac\sigma_b^2$	$\sigma^2 + r\sigma_{abc}^2 + ra\sigma_{bc}^2 + rac\theta_b^2$
C	$\sigma^2 + r\sigma_{abc}^2 + rb\sigma_{ac}^2 + ra\sigma_{bc}^2 + rab\sigma_c^2$	$\sigma^2 + r\sigma_{abc}^2 + rb\sigma_{ac}^2 + ra\sigma_{bc}^2 + rab\sigma_c^2$
AB	$\sigma^2 + r\sigma_{abc}^2 + rc\sigma_{ab}^2$	$\sigma^2 + r\sigma_{abc}^2 + rc\theta_{ab}^2$
AC	$\sigma^2 + r\sigma_{abc}^2 + rb\sigma_{ac}^2$	$\sigma^2 + r\sigma_{abc}^2 + rb\sigma_{ac}^2$
BC	$\sigma^2 + r\sigma_{abc}^2 + ra\sigma_{bc}^2$	$\sigma^2 + r\sigma_{abc}^2 + ra\sigma_{bc}^2$
ABC	$\sigma^2 + r\sigma_{abc}^2$	$\sigma^2 + r\sigma_{abc}^2$
Error	σ^2	σ^2

The difficulties that are encountered in the construction of F_0 statistics for some hypotheses are immediately apparent upon inspection of the expected mean squares. In some instances, there is no legitimate mean square available for the denominator of the F_0 statistic among the existing mean squares. The mean squares for the F_0 statistic must be synthesized as discussed in Sections 5.11 and 7.1.

Alternative Mixed Model with Restrictions on the Interaction

There are several versions of the mixed model based on the definition used for the interaction effects; see Hocking (1973, 1985) and Searle et al. (1992) for a technical discussion. The alternative model places a constraint on the interaction effects. The model is

$$y_{ijk} = \mu + \alpha_i + g_j + (ag)_{ij} + e_{ijk} \qquad (7.7)$$
$$i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, r$$

where μ is the mean, α_i is the fixed effect for factor A, g_j is the random effect for factor B, $(ag)_{ij}$ is the interaction effect, and e_{ijk} is random experimental error. The random effects, g_j and e_{ijk}, are assumed independent and normally distributed with means 0 and variances σ_g^2 and σ^2, respectively. The interaction effects $(ag)_{ij}$ are assumed to be random effects and normally distributed with mean 0 and variance $\frac{(a-1)}{a}\sigma_{ag}^2$.

Since one of the factors, α_i, is fixed the alternative model has the interaction effect sum to zero over levels of the fixed factor, so that

$$\sum_{i=1}^{a} (ag)_{ij} = (ag)_{.j} = 0$$

With this model, interaction effects summed over levels of the random factor $\sum_j^b (ag)_{ij} = (ag)_{i\cdot}$ will not be equal to zero, because they represent only a random sample of the interaction effects at each level of the fixed factor. However, at any given level of the random factor there is a finite set of interaction effects equal to the number of levels for the fixed factor, and the summation, $\sum_i^a (ag)_{ij} = (ag)_{\cdot j}$, is equal to zero. Consequently, there is a covariance between two interaction effects at the same level of the random effect and different levels of the fixed effect, which is $-\frac{1}{a}\sigma_{ag}^2$. For example, with the experiment in Example 7.2 interactions on the same day with the two chemistry methods will be correlated, but interactions for the same chemistry method on two different days will not be correlated.

Thus, the primary distinction for the alternative model is the presence of correlation between the interaction effects. As a consequence, the expectations of some mean squares are different with the alternative model. For example, with two factors, A fixed and B random, the expected mean squares are

$$\begin{aligned} E(MS\boldsymbol{A}) &= \sigma^2 + r\sigma_{ag}^2 + rb\theta_a^2 \\ E(MS\boldsymbol{B}) &= \sigma^2 + ra\sigma_g^2 \\ E[MS(\boldsymbol{AB})] &= \sigma^2 + r\sigma_{ag}^2 \\ E(MSE) &= \sigma^2 \end{aligned} \tag{7.8}$$

The expected mean square for the random main effect factor does not include the interaction component in Equation (7.8), whereas previously in Table 7.5 the interaction component was present. This difference in the expected mean square for the random main effect can have considerable impact on statistical inference. For example, if the restricted model was used for Example 7.2 a test of the hypothesis H_0: $\sigma_d^2 = 0$ would require the test statistic $F_0 = MS\boldsymbol{D}/MSE = 144/14 = 10.29$ and the null hypothesis would be rejected with $F_0 > F_{.05,3,8} = 4.07$, which is just opposite of the conclusion with the original test for the model without restrictions on the interaction.

Hocking (1973) discusses the relationship between the two models and shows the relationship between the variance components for the two models to be

$$\begin{aligned} \sigma_g^2 &= \sigma_b^2 + \frac{1}{a}\sigma_{ab}^2 \\ \sigma_{ag}^2 &= \sigma_{ab}^2 \end{aligned} \tag{7.9}$$

The original model without restrictions on the interaction terms assumed the random interaction effects $(ab)_{ij}$ were uncorrelated with mean 0 and variance σ_{ab}^2. Also, b_i and e_{ijk} are uncorrelated random effects with variances σ_b^2 and σ^2, respectively. There was no assumption that the sum of $(ab)_{ij}$ over levels of the fixed factor $\sum_i^a (ab)_{ij} = (ab)_{\cdot j}$ is equal to zero.

Therefore, a reasonable choice must be made with regard to the model that is most appropriate for the experimental situation. The model without restrictions on the interactions has one major advantage, which is that the expected mean squares for unbalanced data are consistent with the unrestricted model (Hartley & Searle, 1969). The restricted model is not considered in the unbalanced case. If there is a possibility of correlation between effects of a fixed factor for a given level of the random effect, and data are balanced, then the restricted interaction model may be appropriate. If not, or if data are unbalanced, then the model discussed originally in this section with no correlation among the interaction effects is most appropriate.

7.3 Nested Factor Designs: A Variation on the Theme

The standard factorial treatment design has two prominent characteristics: Each level of every factor occurs with all levels of the other factors, and the interaction among factors can be examined.

In certain types of studies the levels of one factor, B, will not be identical across all levels of another factor, A. Each level of factor A will contain different levels of factor B.

The levels of B are said to be nested within the levels of A. The designs are referred to here as **nested factor designs.** They are also called *hierarchical designs*. The following example illustrates a design to study components of variance for nested factors.

Example 7.3 Glucose Standards in Clinical Chemistry

Clinical laboratories perform patient serum assays critical for correct medical diagnoses. The laboratories maintain quality control programs to monitor the performance of assays and ensure the physician is receiving accurate information for diagnosis.

The important sources of variation in the assays are days on which the assays are conducted, the replicate runs within days, and the replicate serum sample preparations within runs. The quality control program requires that a spectrophotometer be tested with several runs on each of several days with serum standards used in the laboratory for control runs. Replicate serum preparations are evaluated within each of the runs.

The data in Table 7.7 are observations from a design used for quality control on the analysis of glucose standards. Glucose standard serums are maintained in the laboratory specifically for quality control runs. There were $c = 3$ replications of the standard prepared for each of $b = 2$ runs on each of $a = 3$ days.

The design is a nested design with two independent and unique runs on each of three days. The nesting of runs within days occurs because a run on any one day has no relationship with a run on any other day. For example, the first run on day 1 has nothing in common with the first runs on days 2 or 3.

The runs are numbered 1 through 6 in Table 7.7 to reflect their independence from one another from day to day.

In the same manner the replicate serum preparations are nested within runs. The Days, Runs, and Replicates represent factors at the first, second, and third levels of the hierarchy.

Table 7.7 Glucose (mg/dl) in quality control standards

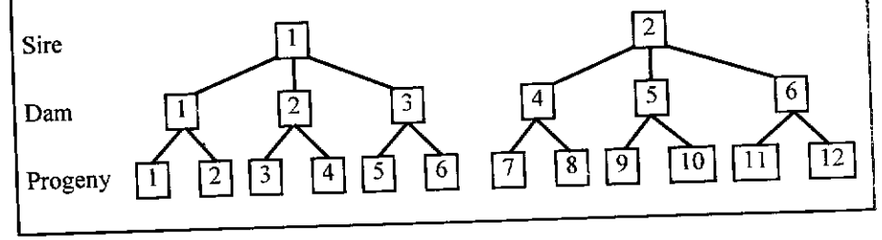
	Day 1		*Day 2*		*Day 3*	
	Run 1	*Run 2*	*Run 3*	*Run 4*	*Run 5*	*Run 6*
	42.5	42.2	48.0	42.0	41.7	40.6
	43.3	41.4	44.6	42.8	43.4	41.8
	42.9	41.8	43.7	42.8	42.5	41.8
Day mean ($\bar{y}_{i..}$)	42.4		44.0		42.0	

Source: Dr. J. Anderson, Beckman Instruments, Inc.

Other Examples with Nested Factors

Consider a genetics study involving animals wherein each sire (male parent) is mated to a random sample of dams (female parents), and each mating results in a litter of several offspring as shown in Display 7.1. The Sires, Dams, and Offspring represent factors of interest in the study. There are different dams for each sire, and the dams are nested within sires. The offspring or progeny from a dam are different from those of other dams, and the progeny are nested within dams. The factors in the nested design form a hierarchy. The hierarchy or nesting of the factors is illustrated in Display 7.1. The uppermost level of the hierarchy represents sires, followed by dams and progeny in the second and third levels, respectively.

Display 7.1 Nested Factor Design in a Genetics Study



There are $b = 3$ different dams mated to each of the $a = 2$ sires for a total of $ab = 6$ dams, and there are $c = 2$ progeny per dam for a total of $abc = 12$ progeny.

Nested designs occur in education research that utilizes several elementary schools. The classrooms are nested within the schools, and the students are nested within the classrooms.

An experiment on fabric dye formulations requires that several replicate batches of each formulation be independently mixed and each batch tested on several specimens of a common fabric. The replicate batches are nested within dye formulations, and fabric specimens are nested within batches.

Statistical Model for Nested Factors

The factors in the nested design hierarchy can be fixed or random factors. The design for glucose standards in Example 7.3 would have all random factors if days, runs, and replicate serum preparations were considered random samples of their respective populations.

The linear model for a nested design with three random nested factors—A, B within A, and C within B—is

$$y_{ijk} = \mu + a_i + b_{j(i)} + c_{k(ij)} \qquad (7.10)$$

$$i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, c$$

where a_i is the effect of factor A, $b_{j(i)}$ is the effect of factor B nested within A, and $c_{k(ij)}$ is the effect of factor C nested within B. The subscript $j(i)$ refers to the factor represented by the j subscript nested within the factor represented by the i subscript. The effects a_i, $b_{j(i)}$, and $c_{k(ij)}$ are assumed to be independent random effects with means 0 and variances σ_a^2, $\sigma_{b(a)}^2$, and $\sigma_{c(b)}^2$, respectively.

Factor A is a fixed effect in the genetics study of Display 7.1 if only two sires are available for the study and the investigator wants to restrict the genetic results to those two sires. The dam effects are random if they represent six dams randomly selected from a potential population of dams. If factor A effects are fixed, then the random effect a_i shown in Equation (7.10) is replaced by the fixed effect notation α_i.

Fixed effects for both factors A and B can occur in the study. Suppose a large metropolitan area fire department wants to evaluate the effect of two different fire station crew rotations on crew efficiency in its six fire districts. One rotation is tested in three of the districts chosen at random, and the other rotation is tested in the remaining three districts. A random sample of crews is tested in each district after the evaluation test period. The rotation factor, A, is fixed because only two types of crew rotations are under consideration. Likewise, the district factor, B, is fixed because only the six districts occurring in the metropolitan area are under consideration. If factors A and B are fixed, then the random effects a_i and $b_{j(i)}$ are replaced by fixed effects α_i and $\beta_{j(i)}$ in Equation (7.10).

Analysis for Random Nested Factors

The objectives of studies that utilize the random effects nested design depend on the subject matter of the study, but estimation of the components of variance and tests of hypotheses about the components are involved. The analysis of variance sums of squares partitions are used for both the estimation and testing procedures.

The computations for the analysis of variance are identical to those for subsampling as shown in Table 5.5. The expectations for the mean squares with all effects random are given in the abbreviated analysis of variance outline shown in Table 7.8.

Table 7.8 Expected mean squares for the analysis of variance of a nested design with three random factors, A, B, and C

Source of Variation	*Degrees of Freedom*	*Mean Square*	*Expected Mean Square*
Total	$abc - 1$		
A	$a - 1$	MSA	$\sigma^2_{c(b)} + c\sigma^2_{b(a)} + bc\sigma^2_a$
B within A	$a(b-1)$	$MS(B/A)$	$\sigma^2_{c(b)} + c\sigma^2_{b(a)}$
C within B	$ab(c-1)$	$MS(C/B)$	$\sigma^2_{c(b)}$

Point Estimators for Variance Components

The analysis of variance estimates of the variance components are found by equating the observed mean squares to the expected mean squares and solving for the components of variance. The three estimators for the analysis in Table 7.8 are

$$\widehat{\sigma}^2_{c(b)} = MS(C/B)$$
$$\widehat{\sigma}^2_{b(a)} = \frac{[MS(B/A) - MS(C/B)]}{c} \qquad (7.11)$$

and

$$\widehat{\sigma}^2_a = \frac{[MSA - MS(B/A)]}{bc}$$

The null hypotheses of interest are H_0: $\sigma^2_a = 0$ if the a_i are random effects and H_0: $\overline{\mu}_{1..} = \overline{\mu}_{2..} = \cdots = \overline{\mu}_{t..}$ for fixed effects with the F_0 statistic $F_0 = MSA/MS(B/A)$. The hypothesis H_0: $\sigma^2_{b(a)} = 0$ is tested with $F_0 = MS(B/A)/MS(C/B)$.

Analysis of Glucose Quality Control Standards

The analysis of variance for the glucose standards in Example 7.3 is shown in Table 7.9. Many computer programs will print a table of expected mean squares and variance component estimates upon request. The coefficient for the Day component (σ^2_a) is $bc = 6$, while the coefficient for the Run/Day component is $c = 3$.

Table 7.9 Analysis of variance for glucose quality control standards

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F Value*	*Pr > F*
Total	17	43.88			
Day	2	13.76	6.88	1.26	.400
Run/Day	3	16.36	5.45	4.75	.021
Rep/Run	12	13.76	1.15		

The estimates for the components of variance are

Reps: $\widehat{\sigma}^2_{c(b)} = MS(\text{Rep/Run}) = 1.15$

Run: $\widehat{\sigma}^2_{b(a)} = \dfrac{[MS(\text{Run/Day}) - MS(\text{Rep/Run})]}{c} = \dfrac{(5.45 - 1.15)}{3} = 1.43$

Day: $\widehat{\sigma}^2_a = \dfrac{[MS\text{Day} - MS(\text{Run/Day})]}{bc} = \dfrac{(6.88 - 5.45)}{6} = 0.24$

The estimate of total variance for a glucose standard analysis is

$$\widehat{\sigma}^2_y = \widehat{\sigma}^2_a + \widehat{\sigma}^2_{b(a)} + \widehat{\sigma}^2_{c(b)} = 0.24 + 1.43 + 1.15 = 2.82$$

with standard deviation $\widehat{\sigma}_y = 1.68$. Approximately 9% of the variation is attributed to day-to-day variation, $\widehat{\sigma}^2_a$; 51% to run-within-day variation, $\widehat{\sigma}^2_{b(a)}$; and 41% to replicates-within-runs variation, $\widehat{\sigma}^2_{c(b)}$. The grand mean of the study was $\overline{y}_{...} = 42.8$, and the percent coefficient of variation for the glucose standards in this set of runs was $\%CV = (100)(\widehat{\sigma}_y/\overline{y}_{...}) = (100)(1.68/42.8) = 3.9\%$.

Standard Errors for Means

The variances for the grand mean for the study $\overline{y}_{...}$ and a day mean $\overline{y}_{i..}$, are

$$\sigma^2_{\overline{y}_{...}} = \frac{\sigma^2_{c(b)}}{abc} + \frac{\sigma^2_{b(a)}}{ab} + \frac{\sigma^2_a}{a} \quad \text{and} \quad \sigma^2_{\overline{y}_{i..}} = \frac{\sigma^2_{c(b)}}{bc} + \frac{\sigma^2_{b(a)}}{b} \qquad (7.12)$$

respectively. The estimates are

$$s^2_{\overline{y}_{...}} = \frac{MS\ \text{Day}}{abc} = \frac{6.88}{(3)(2)(3)} = 0.38$$

and

$$s^2_{\bar{y}_{i..}} = \frac{MS(\text{Run/Day})}{bc} = \frac{5.45}{(2)(3)} = 0.91$$

The standard error estimates are $s_{\bar{y}_{...}} = \sqrt{0.38} = 0.62$ and $s_{\bar{y}_{i..}} = \sqrt{0.91} = 0.95$.

The variances, standard deviations, standard errors, coefficient of variation, and means are all useful statistics for quality control monitoring. Cumulative records of these statistics are often maintained in a laboratory and inspected regularly. If the values deviate from some established norms the analyst is alerted to the fact that the process must be investigated for the source or cause of the deviation.

Tests of Hypotheses About Variances

The test of the null hypothesis for the day component of variance is

$$F_0 = \frac{MS\text{Day}}{MS(\text{Run/Day})} = \frac{6.88}{5.45} = 1.26$$

and it is not significant with $Pr > F = .400$ (Table 7.9). The test of the null hypothesis for the run-within-day component of variance is

$$F_0 = \frac{MS(\text{Run/Day})}{MS(\text{Rep/Run})} = \frac{5.45}{1.15} = 4.75$$

and it is significant with $Pr > F = .021$ (Table 7.9). These results indicate the analyst could most effectively increase precision by concentrating efforts to reduce the variation among runs on the same day and replicates within runs to obtain more precise estimates.

Analysis with Fixed-Factor Effects

The F_0 ratios for tests of hypotheses with fixed effects in the nested design can be determined from the expected mean squares shown in Table 7.10. The fixed effects are defined such that $\Sigma_i \alpha_i = 0$ and $\Sigma_j \beta_{j(i)} = 0$ for $i = 1, 2, \ldots, a$.

Table 7.10 Expected mean squares for the analysis of variance of a nested design with fixed and mixed effects for A and B

Mean Square	*A and B Fixed*	*A Fixed and B Random*
MSA	$\sigma^2_{c(b)} + bc\Sigma\,\alpha_i^2/(a-1)$	$\sigma^2_{c(b)} + c\sigma^2_{b(a)} + bc\Sigma\,\alpha_i^2/(a-1)$
$MS(B/A)$	$\sigma^2_{c(b)} + c\Sigma\,\beta^2_{j(i)}/a(b-1)$	$\sigma^2_{c(b)} + c\sigma^2_{b(a)}$
$MS(C/B)$	$\sigma^2_{c(b)}$	$\sigma^2_{c(b)}$

With A fixed and B random the null hypothesis for factor A effects. H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_a$, is tested with $F_0 = MSA/MS(B/A)$. When both factors are fixed the null hypothesis for factor A effects is tested with $F_0 = MSA/MS(C/B)$. The null hypothesis for factor B effects with both factors fixed is H_0: $\beta_{1(i)} = \beta_{2(i)} = \cdots = \beta_{b(i)}$ for all i, and it is tested with $F_0 = MS(B/A)/MS(C/B)$.

Staggered Nested Designs to Equalize Information About Variances

The nested factors design contains more information on factors at lower levels in the hierarchy of the design than at higher levels. In larger studies the discrepancies in degrees of freedom among sources of variation can be considerable. *Staggered nested designs* were developed to equalize the degrees of freedom for the mean squares at each level of the hierarchy.

The staggered designs have unequal numbers of levels for factors that are nested within other factors. The levels for factor B nested within factor A vary from one level of factor A to another in such a way that the degrees of freedom for MSA and $MS(B/A)$ are more equal. Likewise, the levels for factor C nested within factor B can vary across levels of factor B to achieve degrees of freedom for $MS(C/B)$ similar to that for the other mean squares.

Anderson (1960) and Bainbridge (1963) gave some of the early results on the use of staggered designs. Smith and Beverly (1981) provide a general discussion on the use and analysis of staggered designs. Goldsmith and Gaylor (1970) enumerated 61 staggered designs for three stages of factors (A, B, and C), such that all three variance components can be estimated by the analysis of variance method. The designs enumerated by Goldsmith and Gaylor included two or three of the five possible fundamental structures in a staggered design shown in Display 7.2.

Display 7.2 Five Fundamental Structures for Three-Stage Staggered Nested Designs (Levels of B within A are b_i and the levels of C within B are c_{ij})

	Structures				
Stage	1	2	3	4	5
A					
B/A					
C/B					
b_i	2	2	2	1	1
c_{ij}	2	2, 1	1	2	1

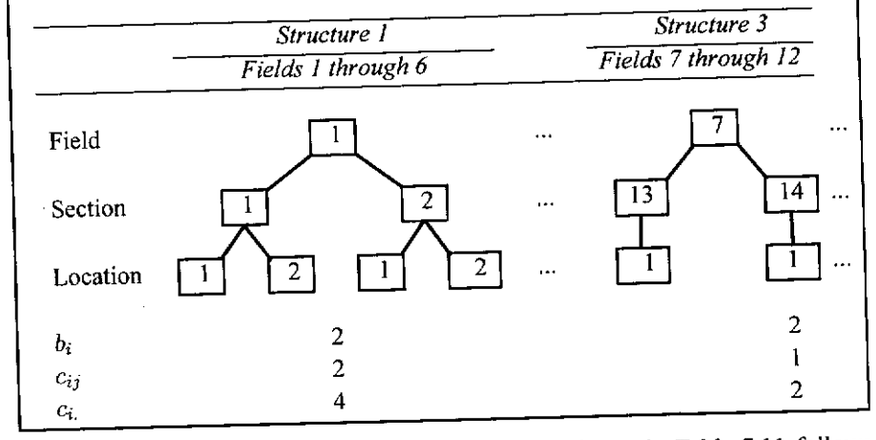
The analysis of a three-stage staggered design follows the pattern established for the analysis with unequal subsamples and replications in Table 5.8. Leone et al. (1968) gave computational formulae for the four-stage nested designs with unequal levels at all stages. They provided formulae for the analysis of variance sums of squares, expected mean squares, and estimates of variances for the means at various stages of the design ($\bar{y}_{....}$, $\bar{y}_{i...}$, $\bar{y}_{ij..}$, and $\bar{y}_{ijk.}$). Gates and Shiue (1962)

provided computational formulae for the analysis of variance sums of squares and expected mean squares for a general S-stage hierarchical classification.

Example 7.4 A Staggered Design for Soil Samples

The data in Exercise 5.8 were generated from a staggered nested design to estimate components of variance for characteristics of soil samples. The three factors in the design were fields (F), sections within fields (S), and locations within sections (L). Suppose $a = 12$ fields and $b_i = 2$ sections per field are sampled. Suppose in fields 1 through 6 that $c_{ij} = 2$ locations per section are sampled, and in fields 7 through 12 $c_{ij} = 1$ location per section is sampled. This staggered design has six replications each of structures 1 and 3 from Display 7.2 as shown in Display 7.3. The design is one of those listed by Goldsmith and Gaylor (1970).

Display 7.3 Staggered Nested Design for Soil Samples



The analysis of variance outline for this design shown in Table 7.11 follows the format shown in Table 5.8 and utilizes Equation (5.23) to compute the coefficients for the variance components in the expected mean squares. The near equal degrees of freedom for the mean squares provide near equal information on each source of variation. The variance components are estimated by the analysis of variance method as

$$\widehat{\sigma}^2_{c(b)} = MS(\boldsymbol{L/S}), \qquad \widehat{\sigma}^2_{b(a)} = \frac{[MS(\boldsymbol{S/F}) - MS(\boldsymbol{L/S})]}{1.50} \tag{7.13}$$

and

$$\widehat{\sigma}^2_a = \frac{[1.50MS\boldsymbol{F} - 1.48MS(\boldsymbol{S/F}) - (1.50 - 1.48)MS(\boldsymbol{L/S})]}{(2.97)(1.50)}$$

Table 7.11 Analysis of variance outline with expected mean squares for the staggered nested design

Source of Variation	*Degrees of Freedom*	*Mean Square*	*Expected Mean Square*
Field	$a - 1 = 11$	MSF	$\sigma^2_{c(b)} + 1.48\sigma^2_{b(a)} + 2.97\sigma^2_a$
Sections	$\sum_{i=1}^{a} b_i - a = 12$	$MS(\boldsymbol{S/F})$	$\sigma^2_{c(b)} + 1.50\sigma^2_{b(a)}$
Locations	$N - \sum_{i=1}^{a} b_i = 12$	$MS(\boldsymbol{L/S})$	$\sigma^2_{c(b)}$

7.4 Nested and Crossed Factors Designs

Certain experimental conditions give rise to factorial arrangements that contain crossed and nested factors. In this case, some factors are crossed in the usual factorial arrangement of levels, whereas other factors are nested within the cells of the factorial arrangement or within levels of at least one other factor. These designs are sometimes called *nested factorial designs* (Anderson & McLean, 1974; Hicks, 1973; Smith & Beverly, 1981).

Example 7.5 Spectrophotometer Evaluation

Companies that manufacture machines and instruments have research and development departments to create new instruments or improve their current line of instruments. Performance tests are a regular part of the development stage of any new machine. The machines are tested for their mechanical or electrical functions, their accuracy in performing their designated function, and so forth.

Research Problem: In one such setting a researcher was developing a new spectrophotometer for applications in medical laboratories. A model spectrophotometer had been constructed according to the proposed design and was ready for evaluation of its spectral capabilities in a laboratory setting. It was necessary to determine whether this particular design determined spectral properties over the required range of serum glucose standards. The researcher had to determine whether the variability and consistency of results over multiple runs and days were within the required specifications.

Treatment Design: A factorial treatment design was used with "concentrations" of glucose and "days" as factors. The serum samples were enhanced with three different levels of glucose to cover the range of glucose concentrations the instrument should be able to analyze. All three concentrations were analyzed on each day, so that concentrations were crossed with

days in a 3×3 factorial arrangement. Two runs of the instrument were made on each day, so that runs were nested within each day.

Experiment Design: Four replicate serum samples were prepared for each of the three concentrations of the glucose standards each day. Two samples of each concentration were assigned randomly to each run of each day. The six samples were analyzed in random order on each run. The same technician prepared the samples and operated the instrument throughout the experiment.

The design had nested and crossed factors with $a = 3$ concentrations crossed with $b = 3$ days with $c = 2$ runs nested within each day and $r = 2$ replicate serum sample preparations for each concentration in each run. The glucose concentrations observed on the spectrophotometer are shown in Table 7.12.

Table 7.12 Glucose concentrations (mg/dl) for three standard concentration samples from two runs of a spectrophotometer on each of three days

	Day 1		*Day 2*		*Day 3*	
Concen.	*Run 1*	*Run 2*	*Run 3*	*Run 4*	*Run 5*	*Run 6*
1	41.2	41.2	39.8	41.5	41.9	45.5
	42.6	41.4	40.3	43.0	42.7	44.7
2	135.7	143.0	132.4	134.4	137.4	141.1
	136.8	143.3	130.3	130.0	135.2	139.1
3	163.2	181.4	173.6	174.9	166.6	175.0
	163.3	180.3	173.9	175.6	165.5	172.0

Source: Dr. J. Anderson, Beckman Instruments, Inc.

Statistical Model for Nested and Crossed Factors

The statistical model for this particular experiment is

$$y_{ijkl} = \mu + \alpha_i + b_j + c_{k(j)} + (ab)_{ij} + (ac)_{ik(j)} + e_{ijkl} \qquad (7.14)$$

$$i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, c \quad l = 1, 2, \ldots, r$$

where α_i is the fixed effect for Concentration, b_j is the random effect for Day, $c_{k(j)}$ is the random effect for Runs nested within Day, $(ab)_{ij}$ is the random effect for Concentration by Day interaction, $(ac)_{ik(j)}$ is the random effect for Concentration by Run interaction nested within Day, and e_{ijkl} is the random experimental error. The assumptions for the effects are consistent with those indicated in previous sections on random, mixed, and nested models.

The model effects for the two crossed factors, Concentration and Day, follow the usual convention with main effects and interaction. The model effect for Runs within Day follows the usual convention for one factor nested within another. The interaction effect for Concentration by Run nested within Day is the one new feature in the model. Since each concentration is evaluated on each run the two factors

constitute a complete factorial arrangement on each day. Therefore, the Concentration by Run interaction can be evaluated on each day and can be nested or pooled over days since the Runs nested within Day are unique runs on each day.

Expected Mean Squares

The model shown in Equation (7.14) is a mixed model with one of the crossed factors random and the other fixed. The expected mean squares for the analysis of variance are affected by the model assumed for the crossed factors and the manner by which the other factors are nested in the experiment. Consequently, a variety of expected mean square patterns are possible. The expected mean squares for models with A and B fixed or with A fixed and B random are shown in Table 7.13.

Table 7.13 Expected mean squares for a nested factorial with A and B crossed and C nested within B

Source of Variation	*Degrees of Freedom*	*Expected Mean Square*	
		A Fixed, B and C Random	*A and B Fixed, C Random*
Total	$abcr - 1$		
A	$a - 1$	$\sigma^2 + r\sigma^2_{ac(b)} + cr\sigma^2_{ab} + bcr\theta^2_a$	$\sigma^2 + r\sigma^2_{ac(b)} + bcr\theta^2_a$
B	$b - 1$	$\sigma^2 + r\sigma^2_{ac(b)} + ar\sigma^2_{c(b)} + cr\sigma^2_{ab} + acr\sigma^2_b$	$\sigma^2 + r\sigma^2_{ac(b)} + ar\sigma^2_{c(b)} + acr\theta^2_b$
AB	$(a-1)(b-1)$	$\sigma^2 + r\sigma^2_{ac(b)} + cr\sigma^2_{ab}$	$\sigma^2 + r\sigma^2_{ac(b)} + cr\theta^2_{ab}$
C/B	$b(c-1)$	$\sigma^2 + r\sigma^2_{ac(b)} + ar\sigma^2_{c(b)}$	$\sigma^2 + r\sigma^2_{ac(b)} + ar\sigma^2_{c(b)}$
AC/B	$b(a-1)(c-1)$	$\sigma^2 + r\sigma^2_{ac(b)}$	$\sigma^2 + r\sigma^2_{ac(b)}$
Error	$abc(r-1)$	σ^2	σ^2

Degrees of Freedom

The degrees of freedom for the sources of variation in the analysis of variance follow the usual conventions for the crossed factors, A and B (Section 6.4). The degrees of freedom for C nested within B follow the convention for one factor nested within another (Section 7.3). The degrees of freedom for the AC interaction nested within B follows the nesting degrees of freedom convention. Only in this case the interaction measured at each level of B has $(a-1)(c-1)$ degrees of freedom; therefore, when nested over b levels of B there are $b(a-1)(c-1)$ degrees of freedom. The degrees of freedom for error are the degrees of freedom for each cell, $r - 1$, pooled over the abc cells in the experiment.

Analysis for Nested and Crossed Factors

The analysis of variance for the observations from the spectrophotometer evaluation of Example 7.5 is shown in Table 7.14. Four different mean squares are required as denominators for the F_0 statistics to test hypotheses about the effects or

variance components in the model with one fixed factor and two random factors (see Tables 7.13 and 7.14). The mean square for Error is the denominator of F_0 to test concentration × runs within day, CR/D. The mean square for CR/D is the denominator of F_0 required to test runs within day, R/D, and concentration × day interaction, CD. The mean square CD is the denominator of F_0 required to test differences among the concentration means, C. Synthesized mean squares for F_0 required to test day variation, D, are

$$MSN = MS(D) + MS(CR/D) = 42.4$$

for the numerator and

$$MS = MS(CD) + MS(R/D) = 131.8$$

for the denominator. Obviously, $F_0 = MSN/MS = 42.4/131.8 = 0.32$ will not be significant. Many programs automatically use the mean square listed for experimental error for all of the F_0 statistics, which is not always the correct denominator mean square. Those programs must be given special instructions to compute the correct F_0 statistics if they have the capability.

Table 7.14 Analysis of variance for spectrophotometer glucose measurements from a factorial with nested and crossed factors

Source of† *Variation*	*Degrees of Freedom*	*Sum of Squares*	*Mean Squares*	*F*	*Pr > F*
Total	35	108,934.1			
C	2	108,263.6	54,131.8	1,227.48	.000
D	2	24.9	12.4	*	*
CD	4	176.4	44.1	1.47	.321
R/D	3	263.1	87.7	2.92	.122
CR/D	6	180.2	30.0	21.43	.000
Error	18	25.8	1.4		

*Test with synthesized Mean Square

†C = Concentration, D = Day, R/D = Run nested in Days

As expected the concentration differences were significant, $F_0 = 1227.48$ with $Pr > F = .000$. The concentration × day interaction was not significant, $Pr > F = .321$, which indicates relatively consistent performance of the instrument from day to day with respect to concentration measurements. However, concentration × run within day was significant, $F_0 = 21.43$ and $Pr > F = .000$. Runs within days with $F_0 = 2.92$ and $Pr > F = .122$ was not significant. The run-to-run consistency of the instrument across concentrations requires some inspection. The inconsistency could be due to the operation of the instrument or the lack of consistency in preparation of the samples for each of the concentrations from run to run.

7.5 How Many Replications?

Random Models

Replication numbers to detect desired significant contributions from a component of variance require a value for the λ constant where the F_0 statistic has the central F_{ν_1,ν_2} distribution multiplied by λ^2 (see Section 5.8). The value for λ^2 may be evaluated in general as follows. Let F_0 be the ratio of mean squares $F_0 = MSN/MSD$, where MSN and MSD, respectively, designate the mean squares for the numerator and denominator of F_0. The constant λ^2 is the ratio of the expected mean squares, or $\lambda^2 = E(MSN)/E(MSD)$. Technical details may be found in Graybill (1961). The charts in Appendix Table X are used as described in Section 5.8.

Consider the two-factor random effects model in Section 7.1. A test of H_0: $\sigma_a^2 = 0$ requires $F_0 = MSA/MS(AB)$, so that

$$\lambda^2 = \frac{E(MSA)}{E[MS(AB)]} = \frac{\sigma^2 + r\sigma_{ab}^2 + rb\sigma_a^2}{\sigma^2 + r\sigma_{ab}^2} = 1 + \frac{rb\sigma_a^2}{\sigma^2 + r\sigma_{ab}^2} \qquad (7.15)$$

Mixed Models

The detection of prescribed fixed-factor effects, say factor A, for the two-factor mixed-model experiment requires

$$\Phi^2 = \frac{br\sum_{i=1}^{a}\alpha_i^2}{a(\sigma^2 + r\sigma_{ab}^2)} \qquad (7.16)$$

for the charts in Appendix Table IX for fixed effects. The value of the constant λ for tests about σ_b^2 and σ_{ab}^2 can be determined as shown for the random model.

7.6 Expected Mean Square Rules

The rules for expected mean squares given in this section apply to most balanced designs with equal replication numbers. The number of levels for any factor do not vary within the balanced designs. The designs include crossed factorials, nested factorials, and mixtures of crossed and nested factors. The rules are adapted from those given in various publications, including Bennett and Franklin (1954) and Mason, Gunst, and Hess (1989). Many computer programs have commands to produce expected mean squares for an analysis of variance.

Rules Illustrated with the Unrestricted Mixed Model

The rules are exemplified with a two-factor mixed-model experiment, A fixed and B random, with r replications of each treatment combination.

1. Write out the linear model for the design:

$$y_{ijk} = \mu + \alpha_i + b_j + (ab)_{ij} + e_{k(ij)}$$

$$i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, r$$

Note the replication subscript k is nested within the ijth treatment combination.

2. Construct a two-way table with
   (a) a row for each term in the model, excluding μ, labeled with the model term and
   (b) a column for each subscript used in the model.

3. Over each column subscript write the number of factor levels for the subscript and write "R" if the factor is random and "F" if the factor is fixed.

4. Add another column with entries as the appropriate fixed or random variance component for the effect represented by that row in the table.

		F	R	R	
		a	b	r	
Source		i	j	k	*Component*
A	α_i				θ_a^2
B	b_j				σ_b^2
AB	$(ab)_{ij}$				σ_{ab}^2
Error	$e_{k(ij)}$				σ^2

5. For each row, if the column subscript does not appear in the row effect, enter the number of levels corresponding to the subscript.

		F	R	R	
		a	b	r	
Source		i	j	k	*Component*
A	α_i		b	r	θ_a^2
B	b_j	a		r	σ_b^2
AB	$(ab)_{ij}$			r	σ_{ab}^2
Error	$e_{k(ij)}$				σ^2

6. If a subscript is bracketed in a row effect, place a 1 in cells under those subscripts that are inside the brackets.

		F	R	R	
		a	b	r	
Source		i	j	k	*Component*
A	α_i		b	r	θ_a^2
B	b_j	a		r	σ_b^2
AB	$(ab)_{ij}$			r	σ_{ab}^2
Error	$e_{k(ij)}$	1	1		σ^2

7. (a) For each row, if any row subscript matches the column subscript, enter a 0 if the column represents a fixed factor F and there is a fixed component of variance for the effect represented by the row.
   (b) Enter a 1 in the remaining cells.

		F	R	R	
		a	b	r	
Source		i	j	k	*Component*
A	α_i	0	b	r	θ_a^2
B	b_j	a	1	r	σ_b^2
AB	$(ab)_{ij}$	1	1	r	σ_{ab}^2
Error	$e_{k(ij)}$	1	1	1	σ^2

8. To determine the expected mean square for a specific source of variation:
   (a) Include σ^2 with a coefficient of 1 in all expected mean squares.
   (b) Of the remaining variance components include only those whose corresponding model terms include the subscripts of the effect under consideration. For $E(MSB)$ the b_j effect, include σ_{ab}^2 and σ_b^2 in addition to σ^2.
   (c) Cover the columns containing non-bracketed subscripts for the effect under consideration. For α_i cover i and for $e_{k(ij)}$ cover k.
   (d) The coefficient for each component in the $E(MS)$ is the product of the remaining columns of the row for that effect. For $E(MSB)$ the column with j is covered so that only the values in columns i and k are visible. For the $(ab)_{ij}$ row the visible values are 1 and r so that the coefficient for σ_{ab}^2 is $1 \cdot r = r$. For the b_j row the visible values are a and r so that the coefficient for σ_b^2 is $a \cdot r$.

		F	R	R		
		a	b	r		
Source		i	j	k	*Component*	$E(MS)$
A	α_i	0	b	r	θ_a^2	$\sigma^2 + r\sigma_{ab}^2 + br\theta_a^2$
B	b_j	a	1	r	σ_b^2	$\sigma^2 + r\sigma_{ab}^2 + ar\sigma_b^2$
AB	$(ab)_{ij}$	1	1	r	σ_{ab}^2	$\sigma^2 + r\sigma_{ab}^2$
Error	$e_{k(ij)}$	1	1	1	σ^2	σ^2

Illustration: The complete table for expected mean square determination of a mixed-model factorial with crossed and nested factors follows. Factors A and B are fixed and crossed, and factor C is random and nested within B across A. The model is

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + c_{k(j)} + (ac)_{ik(j)} + e_{l(ijk)}$$

Source		F a i	F b j	R c k	R r l	Component	$E(MS)$
A	α_i	0	b	c	r	θ_a^2	$\sigma^2 + r\sigma^2_{ac(b)} + bcr\theta_a^2$
B	β_j	a	0	c	r	θ_b^2	$\sigma^2 + r\sigma^2_{ac(b)} + ar\sigma^2_{c(b)} + acr\theta_b^2$
AB	$(\alpha\beta)_{ij}$	0	0	c	r	θ_{ab}^2	$\sigma^2 + r\sigma^2_{ac(b)} + cr\theta_{ab}^2$
C/B	$c_{k(j)}$	a	1	1	r	$\sigma^2_{c(b)}$	$\sigma^2 + r\sigma^2_{ac(b)} + ar\sigma^2_{c(b)}$
$(AC)/B$	$(ac)_{ik(j)}$	1	1	1	r	$\sigma^2_{ac(b)}$	$\sigma^2 + r\sigma^2_{ac(b)}$
Error	$e_{l(ijk)}$	1	1	1	1	σ^2	σ^2

Alteration of the Rules for Restricted Mixed Models

One alteration of the method in Step 7 is required if the restricted mixed model is used—that is, the model in which the interaction effects are correlated and summation of interaction effects over the fixed effects subscripts is restricted to zero. Step 7 (a) will be "*For each row, if any row subscript matches the column subscript, enter a* 0 *if the column subscript represents a fixed factor* F." Step 7 (b) remains unchanged.

Other Estimation Methods for Variance Components

Only analysis of variance estimators have been considered for estimation of variance components in Chapters 5 and 7. Estimation of variance components by the analysis of variance method is relatively straightforward with balanced data (that is, all data cells contain the same number of observations). Estimation of variance components is much more difficult when data are unbalanced. The first major breakthrough for variance component estimation was made by Henderson (1953), who presented three different adaptations of the analysis of variance method for estimating variance components with unbalanced data and random or mixed models.

Since that time, other methods of estimation have been developed, including maximum likelihood (ML) estimation by Hartley and Rao (1967) and a modification of maximum likelihood known as restricted maximum likelihood (REML) attributed to work by Thompson (1962) and Patterson and Thompson (1971). The MINQUE method to find minimum variance quadratic unbiased estimators can be traced to a variety of authors. These methods of estimation are usually available in many of the more comprehensive statistical packages. A comprehensive treatment of variance component models and estimation can be found in Searle et al. (1992).

EXERCISES FOR CHAPTER 7

1. Cholesterol was measured in the serum samples of five randomly selected patients from a pool of patients. Two independent replicate tubes were prepared for each patient for each of four runs on a spectrophotometer. The objective of the study was to determine whether the relative cholesterol measurements for patients were consistent from run to run in the clinic. The data are mg/dl of cholesterol in the replicate samples from each patient on each run.

	Patient				
Run	*1*	*2*	*3*	*4*	*5*
1	167.3	186.7	100.0	214.5	148.5
	166.7	184.2	107.9	215.3	148.5
2	179.6	193.8	111.6	228.9	158.6
	175.3	198.9	114.4	220.4	154.7
3	169.4	179.4	105.9	208.2	144.7
	165.9	177.6	104.1	207.1	145.9
4	177.7	190.4	113.4	221.0	156.1
	177.1	192.4	114.6	219.7	151.0

Source: Dr. J. Anderson, Beckman Instruments, Inc.

a. Write a linear model for the experiment assuming patients and runs are random effects, explain the terms, and conduct an analysis of variance for the data.
b. Show the expected mean squares for the analysis of variance.
c. Estimate the components of variance for runs, patients, and interaction.
d. State the null and alternate hypotheses for main effects and interaction, test each of the null hypotheses, and interpret your results.

2. An animal scientist conducted an experiment to study the effect of water quality on feedlot performance of steer calves. Four water quality treatments were used for the experiment. The water sources were designated as normal (N) and saline (S). The saline water was formulated to approximate the mineral concentrations in some underground water sources utilized in practice for watering livestock. Four combinations of water used in two consecutive 56-day periods of the experiment were N-N, N-S, S-N, and S-S. The feeding trial consisted of the four water treatments with two replicate pens of animals for each treatment in a completely randomized design. The trial was conducted on two separate occasions (two consecutive summers). The resulting design is a factorial arrangement of four water treatments and two summers. The water treatments are considered fixed effects and summers are considered random effects, so that a mixed model is appropriate for the study. The data are the average daily gains for the 16 pens of steers.

	Water			
Summer	*N-N*	*N-S*	*S-N*	*S-S*
1	2.65	2.46	2.56	2.43
	2.53	2.36	2.38	2.50
2	2.25	1.95	2.01	2.14
	2.20	2.25	1.98	2.37

Source: Dr. D. Ray, Department of Animal Sciences. University of Arizona.

a. Write a linear model for the experiment, explain the terms, and conduct an analysis of variance for the data.
b. Prepare a table of cell and marginal means and their respective standard errors.
c. Show the expected mean squares for the analysis of variance.
d. Test the null hypotheses for main effects and interaction, and interpret your results.
e. The water treatments have a 2×2 factorial arrangement. The first factor (A) is normal or saline water in the first 56-day period, and the second factor (B) is normal or saline water in the second 56-day period. Write the linear model for the experiment with this arrangement, considering summers as random effects and factors A and B as fixed effects. Repeat parts (a) through (d) of this exercise with the new model.

3. Three formulations of an alloy were prepared with four separate castings for each formulation. Two bars from each casting were tested for tensile strength. The data are tensile strengths of the individual bars. There are four castings nested within each alloy.

	Castings			
Alloys	*1*	*2*	*3*	*4*
A	13.2	15.2	14.8	14.6
	15.5	15.0	14.2	15.1
B	17.1	16.5	16.1	17.4
	16.7	17.3	15.4	16.8
C	14.1	13.2	14.5	13.8
	14.8	13.9	14.7	13.5

a. Write a linear model for the experiment, assuming alloys as fixed effects and castings within alloys and bars within castings as random effects. Explain the terms, and compute the analysis of variance.
b. Show the expected mean squares for the analysis.
c. Test the null hypothesis for alloy effects, and interpret your results.
d. Compute the estimated means and the 95% confidence interval estimates for the means of each alloy.
e. Estimate the components of variance for castings and bars.

4. A traffic engineering study was conducted to evaluate the effects of three traffic signal types on traffic delay at intersections. The study was also designed to evaluate two methods for measuring

traffic delay. The three signal types were pre-timed, semi-actuated, and fully actuated signals. The two methods, point-sample and path-trace, estimated stopped time per vehicle at an intersection.

Two intersections were used for each signal type. Measurements were made during rush hour and nonrush hour periods. The three crossed factors in the study were signal type, method, and time of day (rush hour and nonrush hour). The intersections were nested within signal types but crossed with method and time of day since both methods were used on the same intersection during both times of day. The data were traffic delay measured as seconds per vehicle.

		Point-Sample		*Path-Trace*	
Signal	*Intersection*	*Rush*	*Nonrush*	*Rush*	*Nonrush*
Pretimed	1	61.7	57.4	53.1	36.5
	2	35.8	18.5	35.5	15.9
Semi-actuated	3	20.0	24.6	17.0	21.0
	4	2.7	3.1	1.5	1.1
Fully actuated	5	35.7	26.8	35.4	20.7
	6	24.3	25.9	27.5	23.3

Source: W. Reilly, C. Gardner, J. Kell (1976), A technique for measurement of delay at intersections. *Technical Report*, FHWA–RD–76-135, Federal Highway Administration, Office of R&D, Washington, D.C.

a. Consider only one method of measuring stopped delay (point-sample or path-trace). Write a linear model for the study, assuming signal type and time of day are fixed effects. Sketch the analysis of variance table, including source of variation, degrees of freedom, and expected mean squares.
b. Suppose you suspect an intersection × time-of-day interaction. Will you be able to test a null hypothesis about the interaction?
c. Suppose there is an interaction between intersection and time of day. Which hypotheses can you test from the analysis of variance?
d. Compute the analysis of variance for the point-sample data. State your assumptions about the model, and test the hypotheses that can be tested with your stated model.
e. Now suppose you want to include the factor for method of measurement into the analysis—that is, point-sample versus path-trace as a fixed effect factor. Write a linear model for the analysis, and sketch the analysis of variance table, including sources of variation, degrees of freedom, and expected mean squares.
f. What assumptions are necessary about the interaction between intersection and the other factors for you to test some hypotheses from the analysis of variance?
g. Compute the analysis of variance for the entire data set. State your assumptions about the model, and test the hypotheses that can be tested about model effects.

5. An experiment was conducted to compare the accuracy of two mass spectrometers in measuring the ratio of ^{14}N to ^{15}N. Two soil samples were taken from each of three plots of land treated with ^{15}N. Two subsamples of each sample were analyzed on each of the two machines. The resulting design has machines crossed with plots and samples. However, the samples are nested within plots. The data are ratios of ^{14}N to ^{15}N (after multiplying by 1000).

Plot	*1*		*2*		*3*	
Sample	*1*	*2*	*3*	*4*	*5*	*6*
Machine A	3.833	3.819	3.756	3.882	3.720	3.729
	3.866	3.853	3.757	3.871	3.720	3.768
Machine B	3.932	3.884	3.832	3.917	3.776	3.833
	3.943	3.888	3.829	3.915	3.777	3.827

Source: D. Robinson (1987), Estimation and use of variance components. *The Statistician* 36, 3–14.

a. Write a linear model for the experiment, assuming machines with fixed effects and plots and samples with random effects; explain the terms; and compute the analysis of variance for the data.
b. Show the expected mean squares.
c. Test the null hypothesis of no difference between the means for the two machines.

6. Use the rules given in Section 7.6 to derive the expected mean squares for the following studies or models:
a. the cholesterol study in Exercise 7.1
b. the cattle-feeding trial in Exercise 7.2
c. the alloy casting experiment in Exercise 7.3
d. the traffic study in Exercise 7.4
e. the soils study in Exercise 7.5
f. a four-stage nested design with the model

$$y_{ijkl} = \mu + a_i + b_{j(i)} + c_{k(ij)} + d_{l(ijk)}$$

$$i = 1,2,3,4 \quad j = 1,2,3 \quad k = 1,2 \quad l = 1,2,3$$

g. a model with nested and crossed factors written as

$$y_{ijklm} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + c_{k(ij)} + \delta_l + (\alpha\delta)_{il} + (\beta\delta)_{jl}$$

$$+ (\alpha\beta\delta)_{ijl} + (cd)_{kl(ij)} + e_{m(ijkl)}$$

$$i = 1, 2, \ldots, a \quad j = 1,22, \ldots, b \quad k = 1, 2, \ldots, c \quad l = 1, 2, \ldots, d \quad m = 1, 2, \ldots, r$$

where α_i, β_j, δ_k, and their interactions are fixed effects and $c_{k(ij)}$, $(cd)_{kl(ij)}$, and $e_{m(ijkl)}$ are random effects

7. How would your statistical inference change if the model with restrictions on the interactions had been used for Example 7.5?

8 Complete Block Designs

Experiment designs to improve the precision of results from research studies are the topics of discussion in this chapter and others to follow. Blocking was introduced in Chapter 1 as a method to reduce experimental error variation. Blocking groups the experimental units into homogeneous blocks to compare treatments within a more uniform environment. The designs in this chapter use either one grouping criterion in a randomized complete block design or two grouping criteria in Latin square arrangements. The features, randomization, analysis, and evaluation of these designs are discussed. Extensions of the designs include factorial treatment designs, multiple experimental units per treatment in each block, and subsampling. Conducting analysis when some observations are missing is discussed. The topic for the final section in this chapter is combining the results from several repetitions of the same experiment at several places or several times.

8.1 Blocking to Increase Precision

Our objective is to have precise comparisons among treatments in our research studies. Blocking is a means to reduce and control experimental error variance to achieve more precision.

Previous chapters concentrated on treatment designs and the associated statistical methods for efficient analysis of research hypotheses. All of the illustrations utilized completely randomized designs. However, outside of appropriate experimental unit selection and good research techniques, the completely randomized designs provide no control over experimental error variance. The experimental units are assumed to be relatively homogeneous with respect to the measured response variable in completely randomized designs. However, sometimes sufficient numbers of homogeneous units do not exist for a complete experiment with these designs.