18 Nested Models

18.1 Introduction

A factor is said to be *nested* within a second factor if each of its levels is observed in conjunction with just one level of the second factor. An example can be obtained from the clean wool experiment that was discussed in the last chapter. There, the objective of the experiment was to examine the variability of the "clean content" among bales of wool in a large shipment. Several bales were selected for examination, and several cores were taken from each bale and measured. Each core was taken from only one bale, so the cores (levels of the first factor) are observed in conjunction with only one bale (level of the second factor). In the above language, the cores are *nested within the bales*. In the original experiment, there was only one observation taken on each core. The variability of the different cores could not, therefore, be distinguished from measurement error, and their effects were not included explicitly in the model. Had there been more than one observation per core, we could have included in the model separate effects due to bales, cores nested within bales, and experimental error.

In this chapter we discuss how to recognize nested factors, how to formulate the associated models, and how to analyze the effects in these models. Many of the analysis techniques are similar to those in the previous chapter.

In the next section we discuss some examples of hypothetical experiments involving nested effects, and possible models to represent the data. In Sect. 18.3, we find the estimable contrasts for fixed-effects nested models and develop tests of hypotheses and confidence intervals for these. The more usual setting where the nested effects are random effects is discussed in Sect. 18.4 and, where possible, we borrow the formulae from the fixed effects setting as we did in Chap. 17. The rules of Chaps. 7 and 17 for finding degrees of freedom, sums of squares and expected mean squares and variance components are then extended to encompass nested models. The analysis of nested models using the SAS and R computer packages is discussed in Sects. 18.5 and 18.6, respectively.

18.2 Examples and Models

Nested factors are usually, but not always, random effects, and they are usually, but not always, blocking factors. In the following examples, we give a selection of different situations involving random effects and suggest some reasonable models to represent the data.

A. Dean et al., *Design and Analysis of Experiments*,
Springer Texts in Statistics, DOI 10.1007/978-3-319-52250-0_18

Example 18.2.1 `Machine head experiment`

Hicks (1956) describes a simple experiment to study the differences in the strain readings (the response) of four different heads on each of five different machines. The heads on each machine were supposedly all doing the same job and should have given rise to similar (nonvariable) readings.

Since each head was observed on only one machine, the heads were "nested within machines," giving twenty heads in total. Four observations were taken on each head. The usual two-way analysis of variance model is not appropriate here, since it would read

$$\begin{gathered} Y_{ijt} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijt} \,, \\ \epsilon_{ijt} \sim N(0, \sigma^2) \,, \\ \epsilon_{ijt}\text{'s are mutually independent,} \\ t = 1, \ldots, 4; \quad i = 1, \ldots, 5; \quad j = 1, \ldots, 4, \end{gathered}$$

where, α_i is the effect of the ith machine, β_j is the effect of the jth head, and $(\alpha\beta)_{ij}$ is the extra effect of observing the ith machine and jth head together. This suggests that every head is observed on every machine, which was not the case. Instead, we need a notation that will clearly indicate the nested nature of the factors. One popular notation, which we shall adopt here, is to replace $\beta_j + (\alpha\beta)_{ij}$ by $\beta_{j(i)}$, where the parentheses indicate that we are looking at the head that happens to be numbered as the *jth head on the ith machine*. The *two-way nested model* is then

$$\begin{gathered} Y_{ijt} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijt} \,, \\ \epsilon_{ijt} \sim N(0, \sigma^2) \,, \\ \epsilon_{ijt}\text{'s are mutually independent,} \\ t = 1, \ldots, 4; \quad i = 1, \ldots, 5; \quad j = 1, \ldots, 4. \end{gathered} \tag{18.2.1}$$

We note in passing that the response Y_{ijt} could also be written as a nested effect $Y_{t(ij)}$, since this represents the tth observation that is specific to the (ij)th machine head. However, since this representation is not crucial to the analysis, we will continue to use the notation Y_{ijt} that we have used so far throughout the book.

One final consideration is whether the machine effects and head effects should be fixed or random. Let us first suppose that the five machines are the only machines of this type in the factory and that they are not due for replacement. The experimenter would then be interested in these five machines specifically, and their effects on the response would be modeled as fixed effects. Let us alternatively suppose that machine heads wear out and are continually being replaced. The experimenter would then be interested in the population of heads from which the particular twenty in the experiment were drawn. Consequently, the nested head effect would be modeled as a random effect. The model would be written as

$$\begin{gathered} Y_{ijt} = \mu + \alpha_i + B_{j(i)} + \epsilon_{ijt} \,, \\ \epsilon_{ijt} \sim N(0, \sigma^2) \,, \quad B_{j(i)} \sim N(0, \sigma^2_{B(A)}) \,, \\ \epsilon_{ijt}\text{'s and } B_{j(i)}\text{'s are all mutually independent,} \\ t = 1, \ldots, 4; \quad i = 1, \ldots, 5; \quad j = 1, \ldots, 4, \end{gathered} \tag{18.2.2}$$

where α_i is the effect of the ith machine, and $\sigma^2_{B(A)}$ is the variance of responses from the population of machine heads that could be fitted on these five machines. Notice that all random variables on the right-hand side of the model are assumed to be mutually independent. □

In the previous example there were two treatment factors, one of whose levels were nested within those of the other. In the following experiment, there are two blocking factors, which are nested one within the other.

Example 18.2.2 `Efficiency experiment`

An experiment was run in 1997 by Carina Dalton, Greg Krzys, Scott O'Dee, and Brad Welch to examine the assertion that "a person works more efficiently when there is no one looking over his or her shoulder." Twelve subjects were recruited for the experiment, and three of these were assigned to each of the four experimenters. Each subject was asked to complete a simple task—crossing through every occurrence of the letter "e" on a page of prose. There were two levels of the treatment factor. Level 1 required the assigned experimenter to look over the subject's shoulder while the task was being completed, and level 2 required the experimenter to be elsewhere in the room absorbed in a book. The response was the time taken to complete the task. Each subject was assigned both treatments, but in a randomized order.

The blocking factor in this experiment was subject. However, the subjects each worked with only one experimenter, and so the subject effects were nested within the experimenter effects.

The subjects were graduate students at The Ohio State University. Although they were not selected according to the rules of a simple random sample, let us suppose that they were a reasonable representation of that population. Let us also suppose that the variation among the techniques of the experimenters, who were also graduate students, was representative of a population of student experimenters. It might also be reasonable to assume that some subjects may be more perturbed than others about an experimenter watching them complete the task. In this case, we might wish to include a subject–treatment interaction in the model. However, there is only one observation per subject per treatment, so the subject–treatment interaction could not be distinguished from the random error.

A second possible model would be to include an experimenter–treatment interaction instead of a subject–treatment interaction. Such an interaction might occur if the actions of the four experimenters were not all identical. In this case the model would be

$$
\begin{gathered}
Y_{hqi} = \mu + E_h + S_{q(h)} + \alpha_i + (\alpha E)_{hi} + \epsilon_{hqi}\ , \\
\epsilon_{hqi} \sim N(0, \sigma^2)\ , \quad S_{q(h)} \sim N(0, \sigma^2_{\mathrm{S(E)}})\ , \quad (\alpha E)_{hi} \sim N(0, \sigma^2_{\mathrm{EA}})\ , \\
\epsilon_{hqi}\text{'s, } E_h\text{'s, } S_{q(h)}\text{'s and } (\alpha E)_{hi}\text{'s are all mutually independent,} \\
h = 1, \ldots, 4; \quad i = 1, 2, 3; \quad i = 1, 2.
\end{gathered}
$$

where E_h is the effect of the hth randomly selected experimenter, $S_{q(h)}$ is the effect of the qth randomly selected subject assigned to the hth experimenter, α_i is the effect of the ith treatment, and $(\alpha E)_{hi}$ is the random effect representing the interaction between the hth experimenter and the ith treatment.

Lastly, we may also wish to include a time-order effect in the model, since the subjects may have been able to complete the task faster on the second occasion just due to familiarity. So we could add the extra term γx_{hqi}, where x_{hqi} is 1 or 2 according to whether the (hq)th subject is assigned treatment i on the first or second occasion. □

18.3 Analysis of Nested Fixed Effects

18.3.1 Least Squares Estimates

Consider first the simplest possible fixed-effects nested model—the two-way nested model (18.2.1) that was suggested for the machine head experiment of Example 18.2.1; that is,

$$\begin{gathered} Y_{ijt} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijt} \,, \\ \epsilon_{ijt} \sim N(0, \sigma^2) \,, \\ \epsilon_{ijt}\text{'s are mutually independent,} \\ t = 1, \dots, r_{ij}; \quad i = 1, \dots, a; \quad j = 1, \dots, b. \end{gathered}$$

The error assumptions are examined in the same way as in Chap. 5 for the one-way analysis of variance model. In any model, the estimable contrasts are functions of the expected values of the response variables (see, for example, Sect. 3.4.1, p. 34). In the present model, $E[Y_{ijt}]$ is equal to

$$E[Y_{ijt}] = \mu + \alpha_i + \beta_{j(i)} \,.$$

If we take an average over the subscripts t and j, we find that a comparison of the levels of A averaged over the levels of B is estimable; that is, we can estimate pairwise comparisons such as

$$\left[\alpha_i + \overline{\beta}_{.(i)}\right] - \left[\alpha_s + \overline{\beta}_{.(s)}\right] ,$$

and we can estimate general contrasts such as

$$\sum_{i=1}^{a} c_i \left[\alpha_i + \overline{\beta}_{.(i)}\right] , \quad \text{with} \sum_{i=1}^{a} c_i = 0 \,.$$

We can also compare the effects of those levels of B that were observed in conjunction with the *same* level of A; that is,

$$\left[\alpha_i + \beta_{j(i)}\right] - \left[\alpha_i + \beta_{u(i)}\right] = \beta_{j(i)} - \beta_{u(i)} \,,$$

or, in general,

$$\sum_{j=1}^{b} d_j \beta_{j(i)} \,, \quad \text{with} \sum_{j=1}^{b} d_j = 0 \,, \quad \text{for any given } i \,.$$

To obtain the least squares estimators of estimable contrasts, we use the method of least squares to find parameter estimates that minimize the sum of squared errors

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} e_{ijt}^2 = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} \left(y_{ijt} - \mu - \alpha_i - \beta_{j(i)}\right)^2 \,.$$

Readers with a knowledge of calculus may verify (see Exercise 7) that the least squares estimate of $\mu + \alpha_i + \beta_{j(i)}$ is $\overline{y}_{ij.}$. Consequently, the least squares estimator of

$$\sum_{i=1}^{a} c_i \left[\alpha_i + \overline{\beta}_{.(i)}\right] \quad \text{is} \quad \sum_{i=1}^{a} c_i \overline{Y}_{i..}$$

with $\Sigma c_i = 0$. The corresponding variance is $\Sigma c_i^2 \sigma^2 / r_{i.}$. Similarly, the least squares estimator of

$$\sum_{j=1}^{b} d_j \beta_{j(i)} \quad \text{is} \quad \sum_{j=1}^{b} d_j \overline{Y}_{ij.} \quad \text{for any } i$$

with $\Sigma d_j = 0$. The corresponding variance is $\Sigma d_j^2 \sigma^2 / r_{ij}$.

All of these formulae can easily be adapted to the case where B has a different number of levels for each level of A by replacing b by b_i.

18.3.2 Estimation of σ^2

The error sum of squares is

$$\begin{aligned} ssE &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} \left(y_{ijt} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_{j(i)}\right)^2 \\ &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} \left(y_{ijt} - \overline{y}_{ij.}\right)^2 \qquad (18.3.3) \\ &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{t=1}^{r_{ij}} y_{ijt}^2 - \sum_{i=1}^{a} \sum_{j=1}^{b} r_{ij} \overline{y}_{ij.}^2 \,. \qquad (18.3.4) \end{aligned}$$

A comparison with the formulae in Sect. 6.4 shows that everything that we have written so far about the fixed-effects two-way nested model could have been deduced from the fixed-effects two-way complete model after replacing $\beta_j + (\alpha\beta)_{ij}$ by $\beta_{j(i)}$. Therefore, we may also deduce that the error mean square, $msE = ssE/(n - v)$, gives an unbiased estimate for σ^2, and the corresponding random variable MSE has a chi-squared distribution with $n - v$ degrees of freedom (where $n = r_{..}$ and $v = ab$).

18.3.3 Confidence Intervals

We may obtain a $100(1 - \alpha)\%$ confidence bound for σ^2 from the information in the previous subsection; that is,

$$\sigma^2 \;\leq\; \frac{ssE}{\chi^2_{n-v,1-\alpha}} \,.$$

The derivation of the bound was explained in Sect. 3.4.6.

Confidence intervals for $\Sigma c_i(\alpha_i + \overline{\beta}_{.(i)})$ and for $\Sigma d_j \beta_{j(i)}$ may be obtained using the relevant methods from Chap. 4 together with the formulae

$$\Sigma c_i \overline{Y}_{i..} \;\pm\; w \sqrt{\sum_i \left(\frac{c_i^2}{r_i}\right) msE}$$

and

$$\Sigma d_j \overline{Y}_{ij.} \;\pm\; w \sqrt{\sum_j \left(\frac{d_j^2}{r_{ij}}\right) msE}\,.$$

18.3.4 Hypothesis Testing

We may obtain a test of the null hypothesis that the levels of B have the same effect on the response within every given level of A, that is,

$$H_0^{B(A)} : \{\beta_{1(i)} = \beta_{2(i)} = \ldots = \beta_{b(i)},\ \text{for every } i = 1, \ldots, a\}\,,$$

against the alternative hypothesis $H_A^{B(A)} : \{H_0^{B(A)}$ is not true$\}$ by comparing the sum of squares for error (18.3.3) in the fixed-effects two-way nested model with the sum of squares for error in the reduced (one-way) model. The reduced model is

$$Y_{ijt} = \mu^* + \alpha_i + \epsilon_{ijt}\,,$$

and the error sum of squares is given by (3.4.4), p. 39, with an extra subscript; that is,

$$ssE_0 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{t=1}^{r_{ij}} (y_{ijt} - \overline{y}_{i..})^2\,.$$

The numerator of the test statistic is then

$$msB(A) = \frac{ssB(A)}{a(b-1)}\,,$$

where the number of degrees of freedom for *B(A)* is obtained as the difference between the error degrees of freedom in the reduced and full models; that is,

$$(n-a) - (n-v) \;=\; v - a \;=\; ab - a \;=\; a(b-1)\,,$$

and where

$$\begin{aligned} ssB(A) &= ssE_0 - ssE \\ &= \sum_i\sum_j\sum_t (y_{ijt} - \overline{y}_{i..})^2 - \sum_i\sum_j\sum_t (y_{ijt} - \overline{y}_{ij.})^2 \\ &= \sum_i\sum_j\sum_t y_{ijt}^2 - \sum_i r_{i.}\overline{y}_{i..}^2 - \sum_i\sum_j\sum_t y_{ijt}^2 + \sum_i\sum_j r_{ij}\overline{y}_{ij.}^2 \\ &= \sum_i\sum_j r_{ij}\overline{y}_{ij.}^2 - \sum_i r_{i.}\overline{y}_{i..}^2 \end{aligned} \qquad (18.3.5)$$

$$= \sum_i \sum_j r_{ij}(\overline{y}_{ij.} - \overline{y}_{i..})^2 \,. \qquad (18.3.6)$$

The decision rule for testing $H_0^{B(A)}$ versus $H_A^{B(A)}$ at significance level α is

$$\text{reject } H_0^{B(A)} \quad \text{if} \quad \frac{ssB(A)/a(b-1)}{ssE/(n-ab)} \; > \; F_{a(b-1),n-ab,\alpha} \,.$$

Similarly, the decision rule for testing

$$H_0^A : \{\alpha_i + \overline{\beta}_{.(i)} \;\text{ all equal}\}$$

against the alternative hypothesis $H_A^A : \{H_0^A \text{ is false}\}$ is

$$\text{reject } H_0^A \text{ if } \frac{ssA/(a-1)}{ssE/(n-ab)} \; > \; F_{a-1,n-ab,\alpha} \,,$$

where

$$ssA = \sum_i r_{i.}(\overline{y}_{i..} - \overline{y}_{...})^2 = \sum_i r_{i.}\overline{y}_{i..}^2 - n\overline{y}_{...}^2 \,.$$

Notice that *ssB(A)* in the two-way nested model is equal to $ssB + ssAB$ in the two-way complete model. Also, the degrees of freedom for *B(A)* in the nested model can be obtained as the sum of the degrees of freedom for B and AB in the complete model; that is,

$$(b-1) + (b-1)(a-1) \; = \; a(b-1) \,.$$

This link between the nested model and the corresponding complete model means that when the sample sizes are equal, we can obtain all the formulae we need from the rules in Chap. 7. This remains true for more complicated models also. For example, if we take the nested model

$$Y_{ijkt} = \mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + \epsilon_{ijkt} \,,$$

we have the following equivalences with the terms of the three-way complete model:

$$\begin{aligned} \beta_{j(i)} &= \beta_j + (\alpha\beta)_{ij} \,, \\ \gamma_{k(ij)} &= \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} \,; \end{aligned}$$

so, for example, the sum of squares for *C(AB)* is

$$\begin{aligned} ssC(AB) &= ssC + ssAC + ssBC + ssABC \\ &= \sum_i \sum_j \sum_k r_{ijk}\overline{y}_{ijk.}^2 - \sum_i \sum_j r_{ij.}\overline{y}_{ij..}^2 \\ &= \sum_i \sum_j \sum_k r_{ijk}(\overline{y}_{ijk.} - \overline{y}_{ij..})^2 \,, \end{aligned}$$

with degrees of freedom

$$(c-1) + (a-1)(c-1) + (b-1)(c-1) + (a-1)(b-1)(c-1) = ab(c-1).$$

As with the crossed model, the degrees of freedom for *C(AB)* give a clue to the subscripts needed in the formula for the sum of squares for *C(AB)*; that is, the degrees of freedom $ab(c-1) = abc - ab$ suggest that the sum of squares for *C(AB)* must contain the terms $\overline{y}_{ijk.}$ and $\overline{y}_{ij..}$, the latter with a minus sign.

To obtain the degrees of freedom corresponding to any effect, we notice that the degrees of freedom for A are the same as in the crossed model; that is, $(a-1)$. The degrees of freedom for *B(A)* are $(b-1)+(a-1)(b-1) = a(b-1)$, and those for *C(AB)* are $ab(c-1)$. Thus we see a pattern. The number of degrees of freedom is the product of the numbers of levels corresponding to the factors in parentheses and one less than the numbers of levels corresponding to the factors not in parentheses. We may now modify rules 1 and 2 in Sect. 7.3 for equal sample sizes listed below. We also include rules 3 and 4 here for easy reference, although these remain the same.

1. Write down the name of the main effect or interaction of interest and the corresponding number of levels and subscripts. Include parentheses to denote nesting of factors.
2. The number of degrees of freedom ν for any effect is the product of the numbers of levels corresponding to the factors in parentheses and one less than the numbers of levels corresponding to the factors not in parentheses.
3. Multiply out the number of degrees of freedom and replace each letter with the corresponding subscripts.
4. The sum of squares for testing the hypothesis that a main effect or an interaction is negligible is obtained as follows. Use each group of subscripts in rule 3 as the subscripts of a term $\overline{y}$, averaging over all subscripts not present and keeping the same signs. Put the resulting estimate in parentheses, square it and sum over all possible subscripts. To expand the parentheses, square each term in the parentheses, keep the same signs, and sum over all possible subscripts.

The other rules remain the same. In particular, confidence intervals for $\sum_{i=1}^{a} c_i \left(\alpha_i + \overline{\beta}_{.(i)}\right)$ and for $\sum_{j=1}^{b} d_j \beta_{j(i)}$ may be calculated using the usual multiple-comparison techniques of Chap. 4.

Example 18.3.1 `Plastic experiment`

Consider the following hypothetical experiment in which a manufacturer of molded plastic wishes to replace a standard ingredient by a cheaper alternative. The two ingredients form the two levels of the treatment factor to be studied. The manufacturing company has factories in three different parts of the country, and since different climates may affect the product differently, the experiment is to take place in each of the three locations. Within each factory, two operators oversee two machines each. The experiment will be run during the usual downtime of the machines.

A possible model for the experiment is

$$\begin{gathered} Y_{ijkut} = \mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + \tau_u + (\tau\gamma)_{uk(ij)} + \epsilon_{ijkut}\,, \\ \epsilon_{ijkut} \sim N(0, \sigma^2)\,, \\ \epsilon_{ijkut}\text{'s are mutually independent,} \\ t = 1, \ldots, r; \quad i = 1, 2, 3; \quad j = 1, 2; \quad k = 1, 2; \quad u = 1, 2; \end{gathered}$$

where α_i is the effect of the ith location, $\beta_{j(i)}$ is the effect of the jth operator at the ith location, $\gamma_{k(ij)}$ is the effect of the kth machine that is looked after by the jth operator at the ith location, τ_u is the effect of the uth treatment, $(\tau\gamma)_{uk(ij)}$ is the interaction effect between the uth treatment and (ijk)th

Table 18.1 Degrees of freedom and sums of squares

Effect	Degrees of freedom	Sum of squares
A	$a-1=2$	$bcdr\Sigma_i\overline{y}^2_{i....} - abcdr\overline{y}^2_{.....}$
$B(A)$	$a(b-1)=3$	$cdr\Sigma_i\Sigma_j\overline{y}^2_{ij...} - bcdr\Sigma_i\overline{y}^2_{i....}$
$C(AB)$	$ab(c-1)=6$	$dr\Sigma_i\Sigma_j\Sigma_k\overline{y}^2_{ijk..} - cdr\Sigma_i\Sigma_j\overline{y}^2_{ij...}$
Trt	$d-1=1$	$abcr\Sigma_u\overline{y}^2_{...u.} - abcdr\overline{y}^2_{.....}$
Trt $\times$ $C(AB)$	$ab(c-1)(d-1)=6$	$r\Sigma_i\Sigma_j\Sigma_k\Sigma_u\overline{y}^2_{ijku.} - dr\Sigma_i\Sigma_j\Sigma_k\overline{y}^2_{ijk..}$ $-cr\Sigma_i\Sigma_j\Sigma_u\overline{y}^2_{ij.u.} + cdr\Sigma_i\Sigma_j\overline{y}^2_{ij...}$
Error	$24r-19$, by subtraction	Obtain by subtraction
Total	$n-1=24r-1$	$\Sigma_i\Sigma_j\Sigma_k\Sigma_u\Sigma_t y^2_{ijkut} - abcdr\overline{y}^2_{.....}$

machine, Y_{ijkut} is the tth response (strength measurement) on the uth treatment and (ijk)th machine, and ϵ_{ijkut} is the corresponding random error, assumed to have a normal distribution with mean 0 and variance σ^2. We also assume that the error variables are mutually independent.

The degrees of freedom and sums of squares for each effect are obtained from rules 1–4 listed above this example and are shown in Table 18.1. Using the formula for confidence intervals in Sect. 18.3.3, we may obtain a confidence interval for $\Sigma_u h_u \tau_u$ as

$$\sum_u h_u \overline{y}_{...u.} \pm w\sqrt{\sum_u \frac{h_u^2}{12r} msE}\,.$$

□

18.4 Analysis of Nested Random Effects

18.4.1 Expected Mean Squares

In Chap. 17 we found that we could modify many of the formulae arising from the fixed-effect crossed models to obtain confidence intervals and hypothesis tests for variance components in the corresponding random-effects models. To find the denominators for the hypothesis tests and to find the estimates for variance components, all we need to do is to calculate the expected values of the mean squares arising from the corresponding fixed-effect models. For equal sample sizes, expected mean squares can be obtained using rule 17 of Sect. 17.8.1, p. 647, exactly as for the random- and mixed-effects crossed models. The rules 18–22 for calculating test ratios and confidence intervals also follow exactly as for the crossed models. We will illustrate these via the model that was suggested for the machine head experiment in Example 18.2.1. A suggested model was

$$\begin{gathered}
Y_{ijt} = \mu + \alpha_i + B_{j(i)} + \epsilon_{ijt}\,, \\
\epsilon_{ijt} \sim N(0,\sigma^2)\,, \quad B_{j(i)} \sim N(0,\sigma^2_{B(A)})\,, \\
\epsilon_{ijt}\text{'s and } B_{j(i)}\text{'s are all mutually independent,} \\
t = 1,\ldots,4; \quad i = 1,\ldots,5; \quad j = 1,\ldots,4,
\end{gathered}$$

Table 18.2 Analysis of variance table for a mixed-effects two-way nested model

Source of variation	Deg. of freedom	Sum of squares	Mean square	Expected mean square
A	$(a-1)$	ssA	msA	$Q(\alpha_i) + r\sigma^2_{B(A)} + \sigma^2$
$B(A)$	$a(b-1)$	$ssB(A)$	$msB(A)$	$r\sigma^2_{B(A)} + \sigma^2$
Error	$ab(r-1)$	ssE	msE	σ^2
Total	$abr-1$	$sstot$		
Formulae for equal sample sizes				
$ssA = br\Sigma_i \overline{y}^2_{i..} - abr\overline{y}^2_{...}$			$ssB(A) = r\Sigma_i\Sigma_j\overline{y}^2_{ij.} - br\Sigma_i\overline{y}^2_{i..}$	
$ssE = \Sigma_i\Sigma_j\Sigma_t y^2_{ijt} - r\Sigma_i\Sigma_j\overline{y}^2_{ij.}$			$sstot = \Sigma_i\Sigma_j\Sigma_t y^2_{ijt} - abr\overline{y}^2_{...}$	

where α_i is the effect of the ith machine, $B_{j(i)}$ is the effect of the jth randomly selected head on the ith machine, and $\sigma^2_{B(A)}$ is the variance of responses from the population of machine heads that could be fitted on these five machines. The error assumptions are examined in the same way as in Chap. 5 for the one-way analysis of variance model. More sophisticated techniques of checking other assumptions on a mixed model with nested effects are discussed by Beckman et al. (1987).

The degrees of freedom and sums of squares for the fixed-effects two-way nested model were calculated in Sect. 18.3. These are listed, for equal sample sizes, in Table 18.2.

We first verify that the fixed-effects mean square for error also provides an unbiased estimate for σ^2 in the mixed effects two-way nested model. From Sect. 18.3.2, we know that $MSE = SSE/(ab(r-1))$, where

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{t=1}^{r} Y^2_{ijt} - r\sum_{i=1}^{a}\sum_{j=1}^{b}\overline{Y}^2_{ij.}\,.$$

For the mixed-effects two-way nested model (18.2.2), we have

$$E[Y_{ijt}] \;=\; E[\overline{Y}_{ij.}] \;=\; E[\overline{Y}_{i..}] \;=\; \mu + \alpha_i \quad \text{and } E[\overline{Y}_{...}] \;=\; \mu + \overline{\alpha}_{.}\,.$$

Also,

$$\text{Var}(Y_{ijt}) \;=\; \sigma^2_{B(A)} + \sigma^2$$

and

$$\text{Var}(\overline{Y}_{ij.}) \;=\; \text{Var}\left(\mu + \alpha_i + B_{j(i)} + \frac{1}{r}\sum_{t=1}^{r}\epsilon_{ijt}\right) \;=\; \sigma^2_{B(A)} + \frac{\sigma^2}{r}\,.$$

So,

$$\begin{aligned} E[SSE] \;&=\; \left[abr(\sigma^2_{B(A)} + \sigma^2) \;+\; br\sum_i(\mu+\alpha_i)^2\right] \\ &\quad - \left[abr\left(\sigma^2_{B(A)} + \frac{\sigma^2}{r}\right) \;+\; br\sum_i(\mu+\alpha_i)^2\right] \\ &=\; ab(r-1)\sigma^2\,. \end{aligned}$$

So, $E[MSE] = \sigma^2$ as required. We also have that

$$\begin{aligned}\text{Var}(\overline{Y}_{i..}) &= \text{Var}\left(\mu + \alpha_i + \frac{1}{b}\sum_{j=1}^{b} B_{j(i)} + \frac{1}{br}\sum_{j=1}^{b}\sum_{t=1}^{r}\epsilon_{ijt}\right) \\ &= \frac{\sigma^2_{B(A)}}{b} + \frac{\sigma^2}{br}\,.\end{aligned}$$

Similarly,

$$\text{Var}(\overline{Y}_{...}) \;=\; \frac{\sigma^2_{B(A)}}{ab} + \frac{\sigma^2}{abr}\,.$$

Consequently, the expected value of the sum of squares for A is

$$\begin{aligned}E[SSA] \;&=\; E\left[br\sum_{i=1}^{a}\overline{Y}^2_{i..} - abr\overline{Y}^2_{...}\right] \\ &=\; \left[abr\left(\frac{\sigma^2_{B(A)}}{b} + \frac{\sigma^2}{br}\right) \;+\; br\sum_i(\mu+\alpha_i)^2\right] \\ &\quad -\; \left[abr\left(\frac{\sigma^2_{B(A)}}{ab} + \frac{\sigma^2}{abr}\right) \;+\; abr\sum_i(\mu+\overline{\alpha}_.)^2\right] \\ &=\; r(a-1)\sigma^2_{B(A)} + (a-1)\sigma^2 + br\sum_i(\alpha_i - \overline{\alpha}_.)^2\,.\end{aligned}$$

Then, since $MSA = SSA/(a-1)$, we have

$$\begin{aligned}E[MSA] \;&=\; \frac{br}{a-1}\sum_i(\alpha_i - \overline{\alpha}_.)^2 + r\sigma^2_{B(A)} + \sigma^2 \\ &=\; Q(\alpha_i) + r\sigma^2_{B(A)} + \sigma^2\,.\end{aligned}$$

Similarly, the expected value of the sum of squares for B nested within A is

$$\begin{aligned}E[SSB(A)] \;&=\; E\left[r\sum_{i=1}^{a}\sum_{j=1}^{b}\overline{Y}^2_{ij.} \;-\; br\sum_{i=1}^{1}\overline{Y}^2_{i..}\right] \\ &=\; \left[abr\left(\sigma^2_{B(A)} + \frac{\sigma^2}{r}\right) \;+\; br\sum_i(\mu+\alpha_i)^2\right] \\ &\quad -\; \left[abr\left(\frac{\sigma^2_{B(A)}}{b} + \frac{\sigma^2}{br}\right) \;+\; br\sum_i(\mu+\alpha_i)^2\right] \\ &=\; ar(b-1)\sigma^2_{B(A)} + a(b-1)\sigma^2\,,\end{aligned}$$

and, since $MSB(A) = SSB(A)/(a(b-1))$, we have

$$E[MSB(A)] = r\sigma^2_{B(A)} + \sigma^2\,.$$

These expected mean squares are listed in the last column of Table 18.2, and we may verify that they can all be obtained from rule 17 of Chap. 17. This rule, which applies also to more complicated mixed-effects nested models, says

17. To obtain the expected mean square for a particular main effect or interaction, first make a note of the subscripts on the term representing that particular effect in the model. Write down variance components for the effect of interest, for the error, and for every interaction whose term in the model includes the noted set of subscripts. Gather up all variance components corresponding to fixed effects into one quadratic form Q. Multiply any remaining variance component except σ^2 by the number of observations taken on each level or combination of levels of the corresponding effect (main effect or interaction). Add up the terms.

18.4.2 Estimation of Variance Components

The rules for obtaining confidence intervals for fixed effects or variance components also remain the same as those in Chap. 17 for non-nested models. Thus, we may obtain a confidence interval for a variance component in a mixed-effects nested model as follows:

19. For a random effect, let $U = \Sigma k_i MS_i$ be the mean square or linear combination of mean squares whose expected value is equal to the variance component corresponding to the random effect. An exact or approximate $100(1-\alpha)\%$ confidence interval for this variance component is

$$\left(\frac{xu}{\chi^2_{x,\alpha/2}} \; , \; \frac{xu}{\chi^2_{x,1-\alpha/2}} \right),$$

where

$$x = \frac{[\Sigma k_i (ms_i)]^2}{\Sigma [k_i (ms_i)]^2 / x_i} \, ,$$

and where u is the observed value of U, ms_i is the observed value of MS_i, and x_i is the number of degrees of freedom corresponding to ms_i.

For example, for the mixed-effects two-way nested model, we may estimate the variability of the response due to the effect of B within A as

$$u = \frac{msB(A) - msE}{r} \, .$$

Then, using rule 19, we can obtain a $100(1-\alpha)\%$ confidence interval for $\sigma^2_{B(A)}$ as

$$\left(\frac{xu}{\chi^2_{x,\alpha/2}} \, , \; \frac{xu}{\chi^2_{x,1-\alpha/2}} \right),$$

where

$$x = \frac{(msB(A) - msE)^2}{\frac{msB(A)^2}{a(b-1)} + \frac{msE^2}{ab(r-1)}} \, .$$

18.4.3 Hypothesis Testing

Hypothesis testing rules are also obtained from the rules in Chap. 17:

18. To obtain the denominator of the test statistic for testing the null hypothesis that a main effect or interaction effect is zero, write down the expected mean square for the effect of interest (see rule 17). Cross out the term that would be zero if the null hypothesis were true. The denominator of the test statistic is the mean square, or linear combination of mean squares, u, whose expected value is equal to the remaining expression.
21. For a fixed effect, the decision rule for testing the hypothesis that the effect is zero is the same as that in rule 8, p. 210, for fixed-effects models except that msE is replaced by the denominator u from rule 18 and the number of error degrees of freedom is replaced by x in rule 19.
22. For a random effect, the decision rule for testing the hypothesis H_0 that the corresponding variance component is zero against the alternative hypothesis that it is not zero is

$$\text{reject } H_0 \text{ if } \frac{ms}{u} > F_{\nu,x,\alpha}\,,$$

where ms is the mean square for the effect of interest and ν the corresponding degrees of freedom, u is the observed value of the denominator as in rule 18, and x is the corresponding degrees of freedom calculated as in rule 19.

For example, using the information in the expected mean squares column of Table 18.2 for the mixed-effects two-way nested model, the decision rule for testing the null hypothesis $H_0^{B(A)} : \{\sigma^2_{B(A)} = 0\}$ of no variability in the effect of B within each level of A against the alternative hypothesis $H_A^{B(A)} : \{\sigma^2_{B(A)} > 0\}$ is

$$\text{reject } H_0^{B(A)} \quad \text{if} \quad \frac{msB(A)}{msE} \; > \; F_{a(b-1),ab(r-1),\alpha}\,, \tag{18.4.7}$$

at chosen significance level α.

To test the hypothesis $H_0^A : \{\alpha_1 = \alpha_2 = \cdots = \alpha_a\}$ that the machine effects are the same averaged over their four heads, the decision rule at significance level α is

$$\text{reject } H_0^A \text{ if } \frac{msA}{msB(A)} > F_{a-1,a(b-1),\alpha}\,. \tag{18.4.8}$$

18.4.4 Some Examples

Example 18.4.1 `Machine head experiment, continued`

The data for the machine head experiment are listed in Table 18.3, and the analysis of variance table is shown in Table 18.4. One can show that the p-value for testing the hypothesis of no machine differences is 0.67, and we would conclude no difference in the effect on strain readings of the five machines. The test of the null hypothesis that the variance $\sigma^2_{B(A)}$ of the population of possible heads fitted to the machines is zero has p-value 0.065. Only if our choice of significance level is greater than this value would we conclude nonzero variability among the heads.

Table 18.3 Data for the machine head experiment

Mach.	Head 1				Head 2				Head 3				Head 4			
1	6	2	0	8	13	3	9	8	1	10	0	6	7	4	7	9
2	10	9	7	12	2	1	1	10	4	1	7	9	0	3	4	1
3	0	0	5	5	10	11	6	7	8	5	0	7	7	2	5	4
4	11	0	6	4	5	10	8	3	1	8	9	4	0	8	6	5
5	1	4	7	9	6	7	0	3	3	0	2	2	3	7	4	0

Source Hicks (1956). Copyright © 1956 American Society for Quality. Reprinted with permission

Table 18.4 Analysis of variance table for the machine head experiment

Effect	d.f.	Sum of squares	Mean square	Expected mean square	Ratio
Machine	4	45.075	11.2688	$Q(\alpha_i) + 4\sigma^2_{B(A)} + \sigma^2$	0.5975
Head (mach.)	15	282.875	18.8583	$4\sigma^2_{B(A)} + \sigma^2$	1.7625
Error	60	642.000	10.7000	σ^2	
Total	79	969.950			

An unbiased estimate of $\sigma^2_{B(A)}$ is given by

$$\frac{msB(A) - msE}{r} = \frac{18.8583 - 10.7000}{4} = 2.0396\,,$$

and since,

$$x = \frac{(2.0396)^2}{\frac{(18.8583/4)^2}{15} + \frac{(10.70/4)^2}{60}} = 2.598\,,$$

a 90% confidence interval for $\sigma^2_{B(A)}$ is given by

$$\left(\frac{(2.598)(2.0396)}{\chi^2_{2.598,.05}}\,,\ \frac{(2.598)(2.0396)}{\chi^2_{2.598,.95}}\right) \approx \left(\frac{5.299}{6.90}\,,\ \frac{5.299}{0.22}\right) = (0.77\,,\ 24.09)$$

measured in squared units of strain. □

Example 18.4.2 Soil experiment

Consider an experiment to compare analyses of soil samples with four treatment factors A, B, C, and D, where

A is "method of analysis" and involves $a = 2$ specifically selected methods.
B is "laboratory" and involves $b = 4$ specifically selected labs.
C is "operator conducting the analysis" and there are $c = 3$ randomly selected operators in each lab.
D is "location from which soil was taken" and involves $d = 3$ randomly selected locations.

Suppose the model is

$$\begin{aligned} Y_{ijkut} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + C_{k(j)} + (\alpha C)_{ik(j)} + D_u \\ &\quad + (\alpha D)_{iu} + (\beta D)_{ju} + (\alpha\beta D)_{iju} + \epsilon_{ijkut}\,, \end{aligned}$$

$$C_{k(j)} \sim N(0, \sigma^2_{C(B)})\,; \quad (\alpha C)_{ik(j)} \sim N(0, \sigma^2_{AC(B)})\,; \quad D_u \sim N(0, \sigma^2_D)\,; \quad (\alpha D)_{iu} \sim N(0, \sigma^2_{AD})$$
$$(\beta D)_{ju} \sim N(0, \sigma^2_{BD})\,; \quad (\alpha\beta D)_{iju} \sim N(0, \sigma^2_{ABD})\,; \quad \epsilon_{ijkut} \sim N(0, \sigma^2)$$
$$i = 1, 2\,; \quad j = 1, 2, 3, 4\,; \quad k = 1, 2, 3\,; \quad u = 1, 2, 3\,; \quad t = 1, 2\,;$$

where α_i is the effect of the ith method of analysis, β_j is the effect of the jth laboratory, and $(\alpha\beta)_{ij}$ is the effect of their interaction; $C_{k(j)}$ is the effect of the kth randomly selected operator in the kth laboratory and $(\alpha C)_{ik(j)}$ is the operator $\times$ analysis method interaction; D_u is the effect of the uth randomly selected location from which the soil was selected and $(\alpha D)_{iu}$, $(\beta D)_{ju}$ and $(\alpha\beta D)_{iju}$ are respectively the interactions of the uth soil location with the ith method of analysis, the uth soil location with the jth laboratory, and the three-factor interaction of the uth soil location, ith method of analysis, and jth laboratory. Two observations are taken on each soil sample via each method of analysis by each operator.

The degrees of freedom, sums of squares, and expected mean squares for this model are obtained using rules 17–21 in Sects. 18.4.1–18.4.3 and are shown in Table 18.5.

The decision rule for testing the null hypothesis $H_0^{ABD} : \{\sigma^2_{ABD} = 0\}$ against the alternative hypothesis $H_A^{ABD} : \{\sigma^2_{ABD} > 0\}$ is given by

$$\text{reject } H_0^{ABD} \text{ if } \frac{msABD}{msE} > F_{6,104,\alpha}\,.$$

If this hypothesis is not rejected, we may wish to examine the AB, AD, and BD interactions. The decision rule for testing the null hypothesis $H_0^{BD} : \{\sigma^2_{BD} = 0\}$ against the alternative hypothesis $H_A^{BD} : \{\sigma^2_{BD} > 0\}$ is given by

$$\text{reject } H_0^{BD} \text{ if } \frac{msBD}{msABD} > F_{6,6,\alpha}\,.$$

The test for the AD interaction is similar, utilizing the test statistic $msAD/msABD$. To obtain a suitable denominator for testing

$$H_0^{AB} : \{(\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..} = 0, \quad \text{for all } i, j\}$$

against the alternative hypothesis that the interaction is not zero, we need the denominator of the test statistic to be an unbiased estimator for

$$6\sigma^2_{AC(B)} + 6\sigma^2_{ABD} + \sigma^2\,.$$

Such an estimator is

$$U = MS(AC(B)) + MS(ABD) - MSE\,.$$

This has approximately a χ^2_x distribution with

Table 18.5 Degrees of freedom, sums of squares, and expected mean squares for the soil experiment

Effect	Degrees of freedom	Expected mean square
A	$a-1=1$	$Q(\alpha,\alpha\beta)+6\sigma^2_{AC(B)}+24\sigma^2_{AD}+6\sigma^2_{ABD}+\sigma^2$
B	$b-1=3$	$Q(\beta,\alpha\beta)+12\sigma^2_{C(B)}+6\sigma^2_{AC(B)}+12\sigma^2_{BD}+6\sigma^2_{ABD}+\sigma^2$
AB	$(a-1)(b-1)=3$	$Q(\alpha\beta)+6\sigma^2_{AC(B)}+6\sigma^2_{ABD}+\sigma^2$
$C(B)$	$b(c-1)=8$	$12\sigma^2_{C(B)}+6\sigma^2_{AC(B)}+\sigma^2$
$AC(B)$	$(a-1)b(c-1)=8$	$6\sigma^2_{AC(B)}+\sigma^2$
D	$d-1=2$	$48\sigma^2_D+24\sigma^2_{AD}+12\sigma^2_{BD}+6\sigma^2_{ABD}+\sigma^2$
AD	$(a-1)(d-1)=2$	$24\sigma^2_{AD}+6\sigma^2_{ABD}+\sigma^2$
BD	$(b-1)(d-1)=6$	$12\sigma^2_{BD}+6\sigma^2_{ABD}+\sigma^2$
ABD	$(a-1)(b-1)(d-1)=6$	$6\sigma^2_{ABD}+\sigma^2$
Error	subtraction $=104$	σ^2
Total	$n-1=143$	

Formulae

$$
\begin{aligned}
ssA &= 72\Sigma_i\overline{y}^2_{i....}-144\overline{y}^2_{.....}\\
ssB &= 36\Sigma_j\overline{y}^2_{.j...}-144\overline{y}^2_{.....}\\
ssAB &= 18\Sigma_i\Sigma_j\overline{y}^2_{ij...}-72\Sigma_i\overline{y}^2_{i....}-36\Sigma_j\overline{y}^2_{.j...}+144\overline{y}^2_{.....}\\
ssC(B) &= 12\Sigma_j\Sigma_k\overline{y}^2_{.jk..}-36\Sigma_j\overline{y}^2_{.j...}\\
ssAC(B) &= 6\Sigma_i\Sigma_j\Sigma_k\overline{y}^2_{ijk..}-18\Sigma_i\Sigma_j\overline{y}^2_{ij...}-12\Sigma_j\Sigma_k\overline{y}^2_{.jk..}+36\Sigma_j\overline{y}^2_{.j...}\\
ssD &= 48\Sigma_u\overline{y}^2_{...u.}-144\overline{y}^2_{.....}\\
ssAD &= 24\Sigma_i\Sigma_u\overline{y}^2_{i..u.}-72\Sigma_i\overline{y}^2_{i....}-48\Sigma_u\overline{y}^2_{...u.}+144\overline{y}^2_{.....}\\
ssBD &= 12\Sigma_j\Sigma_u\overline{y}^2_{.j.u.}-36\Sigma_j\overline{y}^2_{.j...}-48\Sigma_u\overline{y}^2_{...u.}+144\overline{y}^2_{.....}\\
ssABD &= 6\Sigma_i\Sigma_j\Sigma_u\overline{y}^2_{ij.u.}-18\Sigma_i\Sigma_j\overline{y}^2_{ij...}-24\Sigma_i\Sigma_u\overline{y}^2_{i..u.}-12\Sigma_j\Sigma_u\overline{y}^2_{.j.u.}\\
&\quad+72\Sigma_i\overline{y}^2_{i....}+36\Sigma_j\overline{y}^2_{.j...}+48\Sigma_u\overline{y}^2_{...u.}-144\overline{y}^2_{.....}\\
sstot &= \Sigma_i\Sigma_j\Sigma_k\Sigma_u\Sigma_t y^2_{ijkut}-144\overline{y}^2_{.....}
\end{aligned}
$$

ssE is obtained by subtraction

$$x=\frac{[MS(AC(B))+MS(ABD)-MSE]^2}{\frac{MSAC(B)^2}{8}+\frac{MS(ABD)^2}{6}+\frac{MSE^2}{104}}\,.$$

Thus the decision rule for testing H_0^{AB} against H_A^{AB} is

$$\text{reject }\;H_0^{AB}\;\text{ if }\;\frac{msAB}{U}\;>\;F_{3,x,\alpha}\,.$$

An unbiased estimate of σ^2_{BD} is

$$U=\frac{msBD-msABD}{12}\,.$$

This has approximately a χ^2_x distribution, where

$$x = \frac{u^2}{\frac{(msBD/12)^2}{6} + \frac{(msABD/12)^2}{6}} ,$$

and an approximate 95% confidence interval for σ^2_{BD} is

$$\left(u/\chi^2_{x,\alpha/2}, \;\; u/\chi^2_{x,1-\alpha/2} \right) .$$

Any one of the main effects A, B, or $C(B)$ can be investigated if the interactions involving the corresponding factor are all negligible. The relevant formulae can be obtained along the same lines as those described above. □

18.5 Using SAS Software

The SAS procedure `PROC GLM` can handle nested effects when they are described in the `MODEL` statement using notation of the form *B(A)*. The `RANDOM` statement is used to obtain expected mean squares. The procedure `PROC MIXED`, which was described briefly in Chap. 17, can also be used. We will illustrate these procedures via the experiment in Sect. 18.5.1.

18.5.1 Voltage Experiment

An experiment was described by David Desmond in the 1954 issue of *Applied Statistics* on reducing the variability of voltage regulators fitted to motor cars. The voltage regulator was required to operate within a range of 15.8–16.4 volts. When the experiment took place, records showed that about 18% of regulators required readjustment during inspection, and sometimes this figure rose to 50%. Despite the inspection procedure, some of the regulators reaching customers were still outside the specification limits, and complaints from customers were considered to be excessive.

The experiment was run in order to measure the variability in the regulator setting operation. Measurements were taken on 64 voltage regulators at each of four testing stations. The 64 regulators were selected at random from several different setting stations. In Table 18.6, we have reproduced the data for six of these setting stations, corresponding to 40 voltage regulators. Since the regulators were selected at random, we model their effects as random effects nested within setting station. For purposes of illustration, we consider the four testing stations and six setting stations as the only stations of interest and model them as fixed effects. In the original article, these were modeled as random effects.

The effect of testing station is crossed with the effect of setting station and with regulator. A model to describe the data can be written as

$$\begin{gathered} Y_{ijk} = \mu + \alpha_i + \beta_j + C_{k(j)} + \epsilon_{ijk} , \\ \epsilon_{ijk} \sim N(0, \sigma^2) , \quad C_{k(j)} \sim N(0, \sigma^2_{C(B)}) , \\ \epsilon_{ijk}\text{'s and } C_{k(j)}\text{'s are all mutually independent} \\ i = 1, \ldots, 4; \quad j = 1, \ldots, 6; \quad k = 1, \ldots, r_j , \end{gathered}$$

where α_i is the effect of the ith testing station, β_j is the effect of the jth setting station, and $C_{k(j)}$ is the effect of the kth randomly selected regulator from the jth setting station.

There is no reason to suspect that the testing stations would differ in their comparative results for different regulators, so there is no reason to expect a regulator × testing station interaction. Since there

Table 18.6 Voltages for the voltage experiment

Set. sta. (*B*)	Regulator (*C*)	Testing station (*A*)				Set. sta. (*B*)	Regulator (*C*)	Testing station (*A*)			
		1	2	3	4			1	2	3	4
1	1	16.5	16.5	16.6	16.6	4	1	16.1	16.0	16.0	16.2
	2	15.8	16.7	16.2	16.3		2	16.5	16.1	16.5	16.7
	3	16.2	16.5	15.8	16.1		3	16.2	17.0	16.4	16.7
	4	16.3	16.5	16.3	16.6		4	15.8	16.1	16.2	16.2
	5	16.2	16.1	16.3	16.5		5	16.2	16.1	16.4	16.2
	6	16.9	17.0	17.0	17.0		6	16.0	16.2	16.2	16.1
	7	16.0	16.2	16.0	16.0		7	16.0	16.0	16.1	16.0
	8	16.0	16.0	16.1	16.0						
2	1	16.0	16.1	16.0	16.1	5	1	15.5	15.6	15.4	15.8
	2	*15.4*	16.4	16.8	16.7		2	15.8	16.2	16.0	16.2
	3	16.1	16.4	16.3	16.3		3	16.2	*15.4*	16.1	16.3
	4	15.9	16.1	16.0	16.0		4	16.2	16.2	16.0	16.1
							5	16.1	16.2	16.3	16.2
							6	16.1	16.1	16.0	16.1
3	1	16.0	16.0	15.9	16.3	6	1	15.5	15.5	15.3	15.6
	2	15.8	16.0	16.3	16.0		2	16.0	15.6	15.7	16.2
	3	15.7	16.2	15.3	15.8		3	16.0	16.4	16.2	16.2
	4	16.2	16.4	16.4	16.6		4	15.8	16.5	16.2	16.2
	5	16.0	16.1	16.0	15.9		5	15.9	16.1	15.9	16.0
	6	16.1	16.1	16.1	16.1		6	15.9	16.1	15.8	15.7
	7	16.1	16.0	16.1	16.0		7	16.0	16.4	16.0	16.0
							8	16.1	16.2	16.2	16.1

Source Desmond (1954). Copyright © 1956 Blackwell Publishers. Reprinted with permission

is only one observation per regulator–testing station combination, we would not be able to distinguish such an interaction from experimental error. A SAS program for analyzing this model is shown in Table 18.7.

PROC GLM

A plot of the standardized residuals (not shown) highlights two rather large outliers. The two outlying observations are those highlighted in italics in Table 18.6, and we notice that they are from different regulators and different testing stations. If these outliers are removed, the output shown in Fig. 18.1 is obtained. The `TEST` option produces the correct denominators for the tests of $H_0^A : \{\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4\}$, $H_0^B : \{\beta_1 = \beta_2 = \cdots = \beta_6\}$ and $H_0^{C(B)} : \{\sigma^2_{C(B)} = 0\}$. If we select an overall significance level of $\alpha = 0.06$ for the three tests and do each test at level $\alpha^* = 0.02$, we see that there is a significant difference between testing stations, but not between setting stations. Also, the variance of the regulators within setting stations appears to be significantly different from zero. Mind you, the F-test for setting stations is somewhat approximate, not only because the denominator is a composite variance estimator, but also because the treatment type III mean squares may be slightly dependent as a consequence of the removal of the two outliers, causing a slight dependence of the numerator and denominator of the F-statistic.

Unbiased estimates of σ^2 and $\sigma^2_{C(B)}$ can be obtained from the listed expected mean squares as $\hat{\sigma}^2 = msE = 0.0268$ and

Table 18.7 SAS program to analyze a mixed-effects nested model

```
DATA VLT;
 * Input setting station (B), regulator (C),
 *       testing station (A), and voltage;
 INPUT B C A VOLTG;
 LINES;
  1 1 1 16.5
  1 1 2 16.5
  : : :   :
  6 8 4 16.1
;
* Plot standardized residuals versus predicted values for all data;
PROC GLM;
  CLASS A B C;
  MODEL VOLTG = A B C(B);
  RANDOM C(B) / TEST;
  LSMEANS A / PDIFF = ALL CL ADJUST = TUKEY;
  LSMEANS B / PDIFF = ALL CL ADJUST = TUKEY E = C(B);
  OUTPUT OUT = RESIDS PREDICTED = PRED RESIDUAL = Z;
PROC STANDARD STD=1.0;
  VAR Z;
PROC PLOT; * or use PROC SGPLOT;
  PLOT Z*PRED = A Z*PRED = B Z*PRED = C / VREF = 0 VPOS = 19 HPOS = 50;
* Analysis without two outliers;
DATA VLT2; SET VLT;
  IF B = 2 AND C = 2 AND A = 1 THEN DELETE;
  IF B = 5 AND C = 3 AND A = 2 THEN DELETE;
PROC GLM;
  CLASS A B C;
  MODEL VOLTG = A B C(B);
  RANDOM C(B) / TEST;
  LSMEANS A / PDIFF = ALL CL ADJUST = TUKEY;
  * The following should be approximately correct;
  LSMEANS B / PDIFF = ALL CL ADJUST = TUKEY E = C(B);
PROC MIXED METHOD = TYPE3;
  CLASS A B C;
  MODEL VOLTG = A B / DDFM = SAT;
  RANDOM C(B);
  LSMEANS A B / CL PDIFF ADJUST = BON;
```

$$\hat{\sigma}^2_{C(B)} = \frac{msC(B) - msE}{3.9499} = \frac{0.2405 - 0.0268}{3.9499} = 0.0541\,,$$

respectively. Thus the variability of the regulator strain readings is estimated to be about twice as large as the experimental error.

A 90% confidence interval for $\sigma^2_{C(B)}/\sigma^2$ can be obtained by adapting the formula (17.3.11) as follows:

$$\frac{1}{c}\left[\frac{msC(B)}{msE\ F_{\nu_1,\nu_2,\alpha/2}} - 1\right] \le \frac{\sigma^2_{C(B)}}{\sigma^2} \le \frac{1}{c}\left[\frac{msC(B)}{msE\ F_{\nu_1,\nu_2,1-\alpha/2}} - 1\right]$$

$$= \frac{1}{c}\left[\frac{0.2405}{0.0268\ F_{34,115,0.05}} - 1\right] \le \frac{\sigma^2_{C(B)}}{\sigma^2} \le \frac{1}{c}\left[\frac{0.2405}{0.0268\ F_{34,115,0.95}} - 1\right].$$

Fig. 18.1 SAS output for the voltage experiment

Results Viewer - sashtml.htm

The GLM Procedure
Dependent Variable: VOLTG

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	42	11.64681161	0.27730504	10.34	<.0001
Error	115	3.08287193	0.02680758		
Corrected Total	157	14.72968354			

The GLM Procedure

Source	Type III Expected Mean Square
A	Var(Error) + Q(A)
B	Var(Error) + 3.9059 Var(C(B)) + Q(B)
C(B)	Var(Error) + 3.9499 Var(C(B))

The GLM Procedure
Tests of Hypotheses for Mixed Model Analysis of Variance
Dependent Variable: VOLTG

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	3	0.583795	0.194598	7.26	0.0002
C(B)	34	8.176383	0.240482	8.97	<.0001
Error: MS(Error)	115	3.082872	0.026808		

Source	DF	Type III SS	Mean Square	F Value	Pr > F
B	5	2.829636	0.565927	2.38	0.0593
Error	34.085	8.115794	0.238102		
Error: 0.9889*MS(C(B)) + 0.0111*MS(Error)					

Since $E[MSC(B)] = \sigma^2 + 3.9499\sigma^2_{C(B)}$, the value of $c =$ is 3.9499. So, using $F_{34,115,0.05} \approx 1.52$ and $F_{34,115,0.95} = (F_{115,34,0.05})^{-1} \approx 1.64^{-1} = 0.61$, the confidence interval becomes

$$1.242 \le \frac{\sigma^2_{C(B)}}{\sigma^2} \le 3.471\,.$$

The general conclusion of the experiment was that the differences between the four testing stations were of little practical importance. However, we note that the residual plots still indicate one or two large residuals, especially from testing station 2, so perhaps testing station 2 should have been examined a little more closely.

Much of the variability in the regulators appeared to be due to the inherent measurement error, and the experimenters concluded that it was not possible to set the regulators within the desired tolerance limits. A quality control scheme to ensure that the current quality did not deteriorate was put in place.

We note in passing that the effect of the outliers on the analysis was actually very small. If the two original outliers had been included in the analysis, the estimates $\hat{\sigma}^2 = 0.0268$ and $\hat{\sigma}^2_{C(B)} = 0.0541$ would have changed to 0.0392 and 0.0461, respectively. There would also be little change in the p-values of the hypothesis tests. There is some benefit in retaining the entire data set, since the coefficient of $\sigma^2_{C(B)}$ in the expected mean squares is then 4.0, as stated by rule 17 on p. 637.

Fig. 18.2 Output from `PROC MIXED` for the voltage experiment

Results Viewer - sashtml.htm

Differences of Least Squares Means

Effect	B	_B	Estimate	Standard Error	DF	Adj Lower	Adj Upper
B	1	2	0.1084	0.1515	34.7	-0.3693	0.5861
B	1	3	0.2839	0.1276	34.3	-0.1185	0.6864
B	1	4	0.1161	0.1276	34.3	-0.2864	0.5185
B	1	5	0.2975	0.1334	34.5	-0.1232	0.7182
B	1	6	0.3594	0.1233	34.3	-0.02945	0.7482
B	2	3	0.1755	0.1550	34.7	-0.3133	0.6644
B	2	4	0.007659	0.1550	34.7	-0.4812	0.4965
B	2	5	0.1891	0.1598	34.8	-0.3149	0.6931
B	2	6	0.2510	0.1515	34.7	-0.2267	0.7286
B	3	4	-0.1679	0.1318	34.3	-0.5835	0.2478
B	3	5	0.01360	0.1374	34.5	-0.4198	0.4470
B	3	6	0.07545	0.1276	34.3	-0.3270	0.4779
B	4	5	0.1815	0.1374	34.5	-0.2519	0.6148
B	4	6	0.2433	0.1276	34.3	-0.1592	0.6458
B	5	6	0.06185	0.1334	34.5	-0.3589	0.4826

PROC MIXED

The model can also be analyzed using the SAS procedure `PROC MIXED` and the analysis of variance approach as in Chap. 17. The SAS statements are shown in Table 18.7. The analysis of variance output from `PROC MIXED` (not shown) would match that generated by `PROC GLM`, because the option `METHOD = TYPE3` implements the same least squares fit and the same analysis based on Type III sums of squares. An advantage of `PROC MIXED` is that it correctly estimates standard errors for means and contrasts. For example, some of the output generated by the `LSMEANS` statement for comparing setting stations (B) is shown in Fig. 18.2. Note that the standard error and associated degrees of freedom depend on the levels compared, due to the data imbalance caused by the removal of the two outliers. Composite variance estimates are used, and the degrees of freedom are obtained via Satterthwaite's approximation, due to the option `DDFM = SAT` in the `MODEL` statement. The changing number of degrees of freedom from one comparison to another indicates that the corresponding variance estimator also changes. Consequently, the Bonferroni method is used, since Tukey's method requires a common variance estimator, though the latter should also be approximately correct for such nearly balanced data. If the data were more than a little imbalanced, it would be preferable to use restricted maximum likelihood in `PROC MIXED` for variance components estimation—an approach to be discussed in Chap. 19.

18.6 Using R Software

The analysis of variance approach can be used in R for the analysis of balanced designs involving random and nested effects. The `aov` function can fit such models. For example, if a model as specified in R includes the terms `A` and `A:B` but not `B`, then `A:B` represents the effects of B nested within A.

Equivalently, the notation `A/B` causes inclusion of the terms `A` and `A:B` if `B` is excluded. Random effects are designated by inclusion of a single `Error` function in the model. For example, if the model includes the terms `A` and `Error(A:B)` but excludes the term `B`, then `A:B` represents random effects of B nested within A. The `aov` function fits models by least squares, the `summary` function provides the corresponding analysis of variance, including the usual (sometimes approximate) F tests for any fixed effects in the model, and the `lsmeans` function implements multiple comparison procedures. This analysis of variance approach using `aov` is appropriate and the computations dependable given a balanced design.

For unbalanced designs involving random effects, the data analysis can be accomplished by alternative methods involving estimation of the variance components by *restricted maximum likelihood* (ReML). This approach can be implemented using the `lmer` function of the `lme4` package to fit the model, the `anova` function to generate tests of fixed effects, and the `lsmeans` function for multiple comparisons. In our programs, we call the `lmerTest` package rather than `lme4`, as the former provides p-values for F-tests of fixed effects.

We will illustrate the above approaches in Sects. 18.6.2 and 18.6.3, respectively, using the experiment introduced in the following section.

18.6.1 Voltage Experiment

An experiment was described by David Desmond in the 1954 issue of *Applied Statistics* on reducing the variability of voltage regulators fitted to motor cars. The voltage regulator was required to operate within a range of 15.8–16.4 volts. When the experiment took place, records showed that about 18% of regulators required readjustment during inspection, and sometimes this figure rose to 50%. Despite the inspection procedure, some of the regulators reaching customers were still outside the specification limits, and complaints from customers were considered to be excessive.

The experiment was run in order to measure the variability in the regulator setting operation. Measurements were taken on 64 voltage regulators at each of four testing stations. The 64 regulators were selected at random from several different setting stations. In Table 18.6 (p. 688), we have reproduced the data for six of these setting stations, corresponding to 40 voltage regulators. Since the regulators were selected at random, we model their effects as random effects nested within setting station. For purposes of illustration, we consider the four testing stations and six setting stations as the only stations of interest and model them as fixed effects. In the original article, these were modeled as random effects.

The effect of testing station is crossed with the effect of setting station and with regulator. A model to describe the data can be written as

$$\begin{gathered} Y_{ijk} = \mu + \alpha_i + \beta_j + C_{k(j)} + \epsilon_{ijk}\,, \\ \epsilon_{ijk} \sim N(0, \sigma^2)\,, \quad C_{k(j)} \sim N(0, \sigma^2_{C(B)})\,, \\ \epsilon_{ijk}\text{'s and } C_{k(j)}\text{'s are all mutually independent} \\ i = 1, \ldots, 4; \quad j = 1, \ldots, 6; \quad k = 1, \ldots, r_j\,, \end{gathered} \tag{18.6.9}$$

where α_i is the effect of the ith testing station, β_j is the effect of the jth setting station, and $C_{k(j)}$ is the effect of the kth randomly selected regulator from the jth setting station.

There is no reason to suspect that the testing stations would differ in their comparative results for different regulators, so there is no reason to expect a regulator × testing station interaction. Since there

Table 18.8 R program and selected output for analysis of a mixed-effects nested model by the analysis of variance approach

```
> voltage.data = read.table("data/voltage.txt", header=T)
> voltage.data = within(voltage.data, {fSetting = factor(Setting);
+                fRegul = factor(Regul); fTesting = factor(Testing) })
> head(voltage.data, 3)
  Setting Regul Testing Voltg fTesting fRegul fSetting
1       1     1       1  16.5        1      1        1
2       1     1       2  16.5        2      1        1
3       1     1       3  16.6        3      1        1

> # Least squares ANOVA
> # Set contrast options for correct lsmeans and contrasts
> options(contrasts = c("contr.sum", "contr.poly"))
> model1 = aov(Voltg ~ fSetting + fTesting + Error(fSetting:fRegul),
+              data=voltage.data)
> summary(model1)

Error: fSetting:fRegul
          Df Sum Sq Mean Sq F value Pr(>F)
fSetting   5   2.91   0.582     2.6  0.043
Residuals 34   7.61   0.224

Error: Within
           Df Sum Sq Mean Sq F value Pr(>F)
fTesting    3   0.70  0.2341    5.97 0.0008
Residuals 117   4.59  0.0392

> # Multiple comparisons: Tukey’s method
> library(lsmeans)
> lsmTesting1 = lsmeans(model1, ~ fTesting)
> summary(contrast(lsmTesting1, method="pairwise", adjust="tukey"),
+         infer=c(T,T), level=0.98, side="two-sided")

 contrast estimate       SE  df  lower.CL  upper.CL t.ratio p.value
 1 - 2     -0.1550 0.044267 117 -0.285431 -0.024569  -3.502  0.0036
 1 - 3     -0.0825 0.044267 117 -0.212931  0.047931  -1.864  0.2494
 1 - 4     -0.1650 0.044267 117 -0.295431 -0.034569  -3.727  0.0017
 2 - 3      0.0725 0.044267 117 -0.057931  0.202931   1.638  0.3616
 2 - 4     -0.0100 0.044267 117 -0.140431  0.120431  -0.226  0.9959
 3 - 4     -0.0825 0.044267 117 -0.212931  0.047931  -1.864  0.2494

Results are averaged over the levels of: fSetting
Confidence level used: 0.98
Conf-level adjustment: tukey method for comparing a family of 4 estimates
P value adjustment: tukey method for comparing a family of 4 estimates
```

is only one observation per regulator–testing station combination, we would not be able to distinguish such an interaction from experimental error.

18.6.2 Analysis Using Least Squares Estimates and `aov`

Table 18.8 contains an R program for analyzing model (18.6.9) using the `aov` function. The `aov` function, which fits models by ordinary least squares and takes an analysis of variance approach to the data analysis, works fine for models including random effects as long as the design is balanced. The `summary` function generates the appropriate F-tests for each fixed effect, including correct denominators for the tests of $H_0^A : \{\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4\}$ and $H_0^B : \{\beta_1 = \beta_2 = \cdots = \beta_6\}$. If we conduct both tests of fixed effects at level $\alpha^* = 0.02$, we see that there is a significant difference between testing stations ($p = 0.0008$) but not between setting stations ($p = 0.043$). Tukey's method is illustrated for comparing the effects of testing station. Code for comparing the effects of setting station is analogous.

Tests for random effects are not generated by the `aov` and `summary` functions. However, for balanced data, the appropriate tests can be constructed by hand from the mean squares and degrees of freedom generated by the `summary` function, based on the corresponding expected mean squares. Using rule 17 for estimation and hypothesis testing (p. 647), one can show that the expect mean square for regulators nested within setting is $\sigma^2 + 4\,\sigma^2_{C(B)}$. So, to test $H_0^{C(B)} : \{\sigma^2_{C(B)} = 0\}$, the appropriate test statistic is $F = msC(B)/msE = 0.224/0.0392 = 5.714$ with 34 and 117 degrees of freedom, the mean square and degrees of freedom values being obtained from Table 18.8. The reader may verify that the null hypothesis would be rejected at level $\alpha^* = 0.02$ for example ($p < 0.001$), corresponding to an overall significance level of $\alpha = 0.06$ for the three tests.

Unbiased estimates of σ^2 and $\sigma^2_{C(B)}$ can be obtained as $\hat{\sigma}^2 = msE = 0.0392$ and

$$\hat{\sigma}^2_{C(B)} = \frac{msC(B) - msE}{4} = \frac{0.224 - 0.0392}{4} = 0.0462\,,$$

respectively. Thus the variability of the regulator strain readings is estimated to be only slightly larger than the experimental error.

A 90% confidence interval for $\sigma^2_{C(B)}/\sigma^2$ can be obtained by adapting the formula (17.3.11) as follows:

$$\frac{1}{c}\left[\frac{msC(B)}{msE\ F_{\nu_1,\nu_2,\alpha/2}} - 1\right] \leq \frac{\sigma^2_{C(B)}}{\sigma^2} \leq \frac{1}{c}\left[\frac{msC(B)}{msE\ F_{v-1,n-v,1-\alpha/2}} - 1\right]$$

$$= \frac{1}{c}\left[\frac{0.224}{0.0462\ F_{34,117,0.05}} - 1\right] \leq \frac{\sigma^2_{C(B)}}{\sigma^2} \leq \frac{1}{c}\left[\frac{0.224}{0.0462\ F_{34,117,0.95}} - 1\right].$$

Since $E[MSC(B)] = \sigma^2 + 4\sigma^2_{C(B)}$, the value of $c =$ is 4. So, using $F_{34,117,0.05} \approx 1.533$ and $F_{34,117,0.95} = (F_{117,34,0.05})^{-1} \approx 1.635^{-1} = 0.6116$, the confidence interval becomes

$$0.541 \leq \frac{\sigma^2_{C(B)}}{\sigma^2} \leq 1.732\,.$$

The general conclusion of the experiment was that the differences between the four testing stations were of little practical importance. However, we note that the residual plots still indicate one or two large residuals, especially from testing station 2, so perhaps testing station 2 should have been examined a little more closely.

Much of the variability in the regulators appeared to be due to the inherent measurement error, and the experimenters concluded that it was not possible to set the regulators within the desired tolerance limits. A quality control scheme to ensure that the current quality did not deteriorate was put in place.

One advantage to the `aov` function and ordinary least squares is that residuals are available for checking model assumptions. A plot of the standardized residuals (not shown) highlights two rather large outliers. The two outlying observations are those highlighted in italics in Table 18.6, and we notice that they are from different regulators and different testing stations. If these outliers were removed, one would need to use different methods for the data analysis, as illustrated in the next section.

18.6.3 Analysis Using Restricted Maximum Likelihood Estimation

Whether or not the design is balanced, model (18.6.9) can also be analyzed using restricted maximum likelihood (ReML) estimation. In particular, the variance components are estimated by restricted maximum likelihood, providing estimates which make the observed data most likely, subject to the restriction that the variance component estimates be non-negative. Given these variance component estimates, estimated generalized least squares estimates are computed for the fixed effects—generalized to take into account the unequal variances of observations as well as their correlations, and estimated since this variance-covariance structure is estimated. This approach will provide the same results as the analysis of variance approach if the design is balanced and all variance component estimates are positive, as is true for the voltage experiment. For further information about this approach, see Sect. 19.8.3.

Table 18.9 contains an R program and selected output for analyzing model (18.6.9) via this ReML-based approach, but excluding the two outliers. After the program reads all of the voltage data into the data set `voltage.data`, a new data set `voltage2.data` is created from it by taking the subset of `voltage.data` that satisfies two conditions that exclude the two outliers. For example, the first condition

```
!(fSetting == 2   &   fRegul == 2   &   fTesting == 1)
```

means not (`!`) to include observations with `fSetting` value 2 and (`&`) `fRegul` value 2 and `fTesting` value 1, thereby excluding the first outlier. The second outlier is similarly excluded by the second condition.

The `lmer` function fits model (18.6.9), estimating the variance components by restricted maximum likelihood estimation and the fixed effects by estimated generalized least squares estimation. In the model, specified as

```
Voltg ~ fSetting + fTesting + (1|fSetting:fRegul),
```

the term `(1|fSetting:fRegul)` causes inclusion of the random effects $C_{k(j)}$—one parameter for each combination of `fSetting` and `fRegul`. The `anova` function generates type 3 F-tests of the fixed effects—namely, for the effects of `fSetting` and `fTesting`. Finally, the `lsmeans` command applies Tukey's method for each fixed-effects factor using a 98% confidence level, though the results are only displayed for testing stations.

We note in passing that the effect of the outliers on the analysis was actually very small. Comparing the results in Table 18.8 with the outliers to those in Table 18.9 without the outliers, there is little change in the p-values of the hypothesis tests, and Tukey's method yields the same significant comparisons. There is some benefit in retaining the entire voltage data set, since the analysis of variance approach can be used in R for models involving random effects if the design is balanced, in which case the analysis of variance approach is statistically efficient and provides a more complete data analysis.

Exercises

1. **Viscosity experiment**

   An experiment was described by Johnson and Leone (1977, p. 744) to determine the viscosity of a polymeric material. The material was divided into two samples. The two samples were each divided

Table 18.9 R program and selected output for analysis of a mixed-effects nested model using restricted maximum likelihood estimation

```
> voltage.data = read.table("data/voltage.txt", header=T)
> voltage.data = within(voltage.data, {fSetting = factor(Setting);
+                 fRegul = factor(Regul); fTesting = factor(Testing) })

> # Drop two outliers, then reanalyze the data
> voltage2.data = subset(voltage.data,
+                     !(fSetting == 2 & fRegul == 2 & fTesting == 1)
+                  &  !(fSetting == 5 & fRegul == 3 & fTesting == 2) )

> # REML
> # install.packages("lmerTest")
> library(lmerTest) # Attaches/masks lmer and lsmeans,
>                     # adding p-values to anova()
> model2 = lmer(Voltg ~ fSetting + fTesting + (1|fSetting:fRegul),
+               data=voltage2.data)
> anova(model2) # F-tests for fixed effects

Analysis of Variance Table of type III  with  Satterthwaite
approximation for degrees of freedom
         Sum Sq Mean Sq NumDF DenDF F.value  Pr(>F)
fSetting  0.312  0.0623     5    34    2.32 0.06419
fTesting  0.590  0.1967     3   115    7.33 0.00015

> # Multiple comparisons
> library(lsmeans)
> lsmTesting2 = lsmeans(model2, ~ fTesting)
> summary(contrast(lsmTesting2, method="pairwise", adjust="tukey"),
+          infer=c(T,T), level=0.98, side="two-sided")

 contrast  estimate       SE     df  lower.CL  upper.CL t.ratio p.value
 1 - 2    -0.149846 0.037228 115.22 -0.259570 -0.040121  -4.025  0.0006
 1 - 3    -0.055856 0.036922 115.10 -0.164681  0.052968  -1.513  0.4332
 1 - 4    -0.138356 0.036922 115.10 -0.247181 -0.029532  -3.747  0.0016
 2 - 3     0.093989 0.036921 115.12 -0.014833  0.202811   2.546  0.0583
 2 - 4     0.011489 0.036921 115.12 -0.097333  0.120311   0.311  0.9895
 3 - 4    -0.082500 0.036618 115.00 -0.190429  0.025429  -2.253  0.1154

Results are averaged over the levels of: fSetting
Confidence level used: 0.98
Conf-level adjustment: tukey method for comparing a family of 4 estimates
P value adjustment: tukey method for comparing a family of 4 estimates

> lsmSetting2 = lsmeans(model2, ~ fSetting)
> summary(contrast(lsmSetting2, method="pairwise", adjust="tukey"),
+          infer=c(T,T), level=0.98, side="two-sided")
```

into ten "aliquots." After preparation of these aliquots, they were divided into two subaliquots and a further step in the preparation made. Finally, each subaliquot was divided into two parts and the final step of the preparation made. The viscosity determinations are listed in Table 18.10.

Table 18.10 Viscosity determinations for the viscosity experiment

Sample	Aliquot	Subaliquot 1		Subaliquot 2	
		Part 1	Part 2	Part 1	Part 2
1	1	59.8	59.4	58.2	63.5
	2	66.6	63.9	61.8	62.0
	3	64.9	68.8	66.3	63.5
	4	62.7	62.2	62.9	62.8
	5	59.5	61.0	54.6	61.5
	6	69.0	69.0	60.6	61.8
	7	64.5	66.8	60.2	57.4
	8	61.6	56.6	64.5	62.3
	9	64.5	61.3	72.7	72.4
	10	65.2	63.9	60.8	61.2
2	1	59.8	61.2	60.0	65.0
	2	65.0	65.8	64.5	64.5
	3	65.0	65.2	65.5	63.5
	4	62.5	61.9	60.9	61.5
	5	59.8	60.9	56.0	57.2
	6	68.8	69.0	62.5	62.0
	7	65.2	65.6	61.0	59.3
	8	59.6	58.5	62.3	61.5
	9	61.0	64.0	73.0	71.7
	10	65.0	64.0	62.0	63.0

Source Johnson and Leone (1977).

(a) Write down a model for the viscosity determinations allowing for variability in the samples, aliquots, subaliquots and parts.
(b) Examine the error assumptions on your model.
(c) Estimate the variances of all the random effects in the model.
(d) Give a set of confidence intervals for the variances of all the random effects in the model at overall significance level 90%. At which step of the preparation is most of the variability introduced?

2. **Sleep experiment**

Sleeping patterns can be classified according to periods of "deep sleep" and of "REM sleep" (rapid eye movement). An experiment is done to see how sleeping tablets and amount of daily activity affect the proportion of REM sleep. Three types of sleeping tablets are to be tested, coded 1, 2, 3 (where type 3 is a placebo).

Twelve subjects are selected at random from a large population and are assigned at random to the levels of A, four to each level. Each subject is assigned an activity level for the day, and the proportion of REM sleep is monitored during that night. The four activity levels are:

B1 = read quietly all day, B2 = walk 10 miles during the day,
B3 = spend the day shopping, B4 = play video games all day.

The experiment continues for four days, so that each subject is observed at each activity level in a random order. The model is assumed to be

$$Y_{hijt} = \mu + S_{h(i)} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{hijt}\ ,$$

where α_i is the effect of the ith sleeping tablet, β_j is the effect of the jth activity level, $(\alpha\beta)_{ij}$ is the effect of their interaction, $S_{h(i)}$ is the effect of the hth random subject assigned to the ith sleeping tablet, and $S_{h(i)} \sim N(0, \sigma^2_{S(A)})$ and $\epsilon_{hijt} \sim N(0, \sigma^2)$ and all $S_{h(i)}$ and ϵ_{hijt} are independent.

(a) Write down the degrees of freedom and expected mean squares for the analysis of variance table.
(b) Explain how to test the null hypothesis $H_0^A : \{\alpha_1 = \alpha_2 = \alpha_3\}$ against the alternative hypothesis that at least two of the α_i differ.
(c) Explain how to test the null hypothesis $H_0^{S(A)} : \{\sigma^2_{S(A)} = 0\}$ against the alternative hypothesis $H_A^{S(A)} : \{\sigma^2_{S(A)} > 0\}$.
(d) Suppose that the null hypothesis

$$H_0^{AB} : \{(\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..}, \quad \text{for all } i, j\}$$

appears to be correct. Which contrasts would be of particular interest to the experimenter? Why? Give formulas that would provide an overall 95% set of confidence intervals for your chosen contrasts. Give reasons for your choice of formula(s).
(e) If the experimenter thought that a day effect would be important, how would you modify the design of the experiment and the model?

3. Consider the model

$$Y_{ijkl} = \mu + \alpha_i + B_{j(i)} + C_{k(ji)} + \delta_l + (\alpha\delta)_{il} + (B\delta)_{lj(i)} + \epsilon_{ijkl}\ ,$$

(a) Calculate the expected mean squares for all effects in the model.
(b) Which ratio would you use to test $H_0 : \{\delta_l + \overline{(\alpha\delta)}_{.l}$ all equal$\}$?
(c) Which ratio would you use to test $H_0 : \sigma^2_A = 0$?

4. **Titanium alloy experiment**

An experiment described by Johnson and Leone (1977, p. 758) was performed by a company to investigate the effects of various factors on the "yield strength" of a particular titanium alloy. The factors investigated were:

A: vendors (4 fixed levels representing suppliers of raw material).
C: bar size (2 fixed levels representing standard sizes of bars of raw material).
B: batch (3 randomly selected levels nested within each combination of levels of A and C).
D: product type (2 fixed levels representing different types of finished product—forgedown and finished-forge blades).

Three observations were taken on each treatment combination. A reasonable model was thought to be

$$\begin{aligned} Y_{ijklt} = \mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij} + B_{k(ij)} + \delta_l \\ + (\alpha\delta)_{il} + (\gamma\delta)_{jl} + (B\delta)_{kl(ij)} + \epsilon_{ijklt}\ , \end{aligned}$$

$$\epsilon_{ijklt} \sim N(0, \sigma^2)\ , \quad B_{k(ij)} \sim N(0, \sigma^2_{B(AC)})\ , \quad (B\delta)_{kl(ij)} \sim N(0, \sigma^2_{BD(AC)})\ ,$$

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

where α_i, γ_j, and δ_l represent the effects of the ith vendor, jth bar size, and lth product type, respectively, and $B_{k(ij)}$ represents the effect of the kth randomly selected batch of the jth bar size made with bar stock from the ith vendor, and random variables on the right-hand side of the model are assumed to be mutually independent.

(a) Write down the degrees of freedom and expected mean squares column of the analysis of variance table.
(b) Give a formula for an approximate 95% confidence interval for $\sigma^2_{B(AC)}$.
(c) How would you test the hypothesis

$$H_0 : \{\text{no differences in yield strength of the titanium alloy can be attributed to the four vendors}\}$$

against the alternative hypothesis $H_A : \{H_0 \text{ is false}\}$?

5. **Titanium alloy experiment, continued**

Suppose that factors C and D are to be investigated further in a followup experiment. Suppose that two new factors P and Q ("heat setting during processing" and "cooling method") are also to be investigated at two levels each. A followup experiment is required with the four factors C, D, P, and Q at two levels each (a 2^4 experiment). Only sixteen observations will be taken, four for each vendor. It is known that the interactions CP, CQ, PQ, CPQ, and $CDPQ$ are likely to be negligible. Also, there was information gained from the previous parts to Exercise 4 to suggest that all interactions of treatment factors with vendor can be assumed negligible.

(a) Divide the 16 treatment combinations into four blocks of size four (one block for each vendor). Show your design explicitly, and indicate what should be randomized.
(b) Write down a suitable model and the degrees of freedom column for the analysis of variance table for your design in part (a).
(c) Before your design in part (a) is run, the management announces that in future, only one vendor will be used by the company. Also, your budget is cut, so that you can take only 8 observations. Thus, you need to design a $\frac{1}{2}$–fraction of a 2^4 experiment. In reviewing the list of negligible interactions above, you discover that two have been omitted. Interactions DP and CDQ are also known to be negligible. Choose a design and list the treatment combinations explicitly. (Hint: Try $I = CPQ$.) State the aliasing scheme and a suitable model. Will there be any problems in interpreting the results of this experiment?

6. **Operator experiment**

An experiment to identify the causes of variability in readings of a spectrometer was described in Exercise 10 of Chap. 7, p. 241. The same authors (Inman et al., *Journal of Quality Technology*, 1992) also described a study to determine how much of the variation in measured manganese concentration in steel was due to operator variation.

Ten steel samples were sliced from a steel billet. Each operator was asked to measure the manganese content of each sample twice. The measurements taken by any one operator were done in a random order on a single day. There were four operators, who were regarded as representative of a large population of potential operators.

(a) Write down a model for this experiment. Indicate clearly which effects are fixed, random, crossed, and nested.

Table 18.11 Manganese concentrations (percentages) for the operator experiment

Sample	Operator 1		2		3		4	
1	0.63	0.60	0.62	0.62	0.60	0.60	0.59	0.61
2	0.64	0.63	0.63	0.64	0.67	0.65	0.62	0.64
3	0.60	0.58	0.60	0.61	0.60	0.60	0.58	0.60
4	0.75	0.74	0.74	0.74	0.74	0.73	0.73	0.76
5	0.71	0.68	0.69	0.70	0.69	0.67	0.68	0.71
6	0.65	0.63	0.62	0.65	0.63	0.64	0.62	0.64
7	0.67	0.64	0.66	0.67	0.65	0.65	0.64	0.66
8	0.65	0.63	0.65	0.64	0.62	0.62	0.60	0.62
9	0.68	0.66	0.67	0.68	0.67	0.67	0.65	0.68
10	0.67	0.64	0.66	0.66	0.65	0.64	0.64	0.66

Source Inman, Ledolter, Lenth, and Niemi (1992). Reprinted with Permission from Journal of Quality Technology © 1992 ASQ, www.asq.org

(b) Write down the degrees of freedom, the sums of squares, and the expected mean squares for each of the sources of variation in your model.

(c) The authors analyzed this experiment using a gamma distribution to model the distribution of the error terms. Using the data in Table 18.11, investigate whether or not the normal distribution could be used (it may be necessary to take a transformation).

(d) If the normal distribution can be used as a reasonable approximation to the error distribution, then analyze the experiment. In particular, obtain estimates of the variances of the random effects and identify the major sources of variation.

7. For the two-way nested fixed-effects model (18.2.1) on p. 672, show that the least squares estimator of $\mu + \alpha_i + \beta_{j(i)}$ is given by $\overline{Y}_{ij.}$.
[Hint: Differentiate the sum of squared errors with respect to μ, α_i ($i = 1, \ldots, a$), and $\beta_{j(i)}$ ($j = 1, \ldots, b$; $i = 1, \ldots, a$), in turn. Set the resulting three sets of normal equations equal to zero. Show that the third set of equations adds to the first equation, and that the ith portion of the third set of equations adds to the ith equation in the second set. Thus, the first and second sets of equations are redundant, and $a + 1$ extra equations must be added to the set.]

8. **Red blood cell experiment**

The trout experiment reported by Gutsell (*Biometrics*, 1951) was described in Exercise 15 of Chap. 3. As part of the same experiment, the red blood cell counts in the blood of brown trout were measured. Fish were put at random into eight troughs of water. Two troughs were assigned to each of the four levels of the treatment factor "sulfamerazine" (0, 5, 10, 15 grams per 100 pounds of fish added to the diet per day). After 42 days, five fish were selected at random from each trough and the red blood cell count from the blood of each fish was measured in two different counting chambers, giving two measurements per fish. The observations reported in Table 18.12, when multiplied by 5000, give the number of red blood cells per cubic millimeter of blood.
A possible model for these data is

$$Y_{ijkt} = \mu + \alpha_i + B_{j(i)} + C_{k(ij)} + \epsilon_{ijkt}\,,$$
$$\epsilon_{ijkt} \sim N(0, \sigma^2)\,, \quad B_{j(i)} \sim N(0, \sigma^2_{B(A)})\,, \quad C_{k(ij)} \sim N(0, \sigma^2_{C(AB)})\,,$$
$$i = 1, 2, 3, 4; \quad j = 1, 2; \quad k = 1, \ldots, 5; \quad t = 1, 2;$$

Table 18.12 Red blood cell counts from brown trout for the red blood cell experiment

Fish	0 gm sulf.				5 gm sulf.			
	Trough 1		Trough 2		Trough 1		Trough 2	
1	213	230	166	157	296	319	310	309
2	253	231	206	185	278	258	241	270
3	195	164	245	250	345	307	272	311
4	193	203	213	181	322	372	254	237
5	191	195	198	169	248	274	266	275
Fish	10 gm sulf.				15 gm sulf.			
	Trough 1		Trough 2		Trough 1		Trough 2	
1	339	322	196	232	278	212	287	280
2	282	285	205	186	275	311	221	243
3	236	262	252	274	186	158	331	309
4	252	209	245	216	301	281	231	244
5	263	296	249	260	223	246	292	295

Source Gutsell (1951).

where α_i is the effect of the ith level of sulfamerazine in the diet, $B_{j(i)}$ is the effect of the jth randomly selected trough assigned to the ith level of sulfamerazine, and C_k is the effect of the kth randomly selected fish from the (ij)th trough, and random variables on the right hand-side of the model are assumed to be mutually independent.

(a) What are the experimental units and observational units in this experiment?
(b) Since the data are counts, examine the assumptions of normally distributed errors and equal error variances by treatment. If the assumptions are not approximately satisfied, is there a transformation that can be used to correct the problem?
(c) Write out the degrees of freedom and the expected mean squares for each term in the model.
(d) Test the hypothesis that sulfamerazine has no effect on the red blood cell counts. Examine the linear and quadratic trends.
(e) If the test in part (d) is rejected, calculate simultaneous 95% confidence intervals for pairwise comparisons in the effects of the sulfamerazine levels.