Section 7.3 Nested Design for Models with Fixed Effects, Mixed Effects and Random Effects

TABLE 25.5 Expected Mean Squares for Balanced Two-Factor ANOVA Models.

Mean Square	df	Fixed ANOVA Model (A and B fixed)	Random ANOVA Model (A and B random)	Mixed ANOVA Model (A fixed, B random)
MSA	$a-1$	$\sigma^2 + nb\frac{\sum \alpha_i^2}{a-1}$	$\sigma^2 + nb\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2$	$\sigma^2 + nb\frac{\sum \alpha_i^2}{a-1} + n\sigma_{\alpha\beta}^2$
MSB	$b-1$	$\sigma^2 + na\frac{\sum \beta_j^2}{b-1}$	$\sigma^2 + na\sigma_\beta^2 + n\sigma_{\alpha\beta}^2$	$\sigma^2 + na\sigma_\beta^2$
$MSAB$	$(a-1)(b-1)$	$\sigma^2 + n\frac{\sum\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)}$	$\sigma^2 + n\sigma_{\alpha\beta}^2$	$\sigma^2 + n\sigma_{\alpha\beta}^2$
MSE	$(n-1)ab$	σ^2	σ^2	σ^2

Example 1

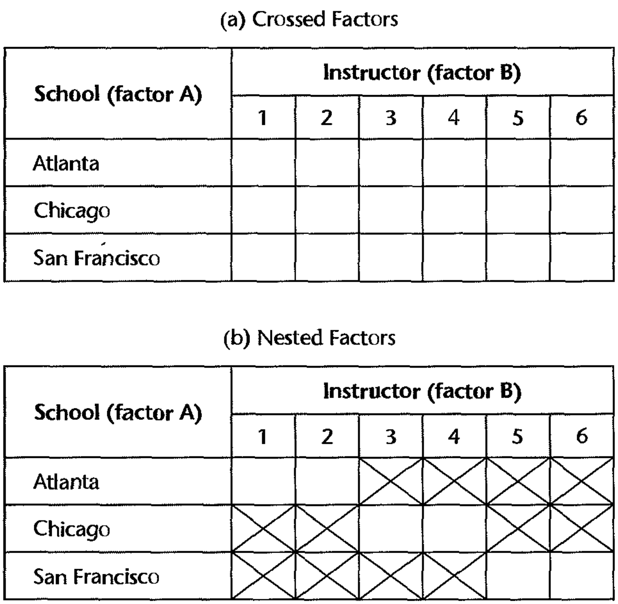
A large manufacturing company operates three regional training schools for mechanics, one in each of its operating districts. The schools have two instructors each, who teach classes of about 15 mechanics in three-week sessions. The company was concerned about the effect of school (factor A) and instructor (factor B) on the learning achieved. To investigate these effects, classes in each district were formed in the usual way and then randomly assigned to one of the two instructors in the school. This was done for two sessions, and at the end of each session a suitable summary measure of learning for the class was obtained. The results are presented in Table 26.1.

The layout of Table 26.1 appears identical to an ordinary two-factor investigation, with two observations per cell (see, e.g., Table 19.7). In fact, however, the study is not an ordinary two-factor study. The reason is that the instructors in the Atlanta school did not also teach in the other two schools, and similarly for the other instructors. Thus, six different instructors were involved. An ordinary two-factor investigation with six different instructors would have consisted of 18 treatments, as shown in Figure 26.1a. In the training school example, however, only six treatments were included, as shown in Figure 26.1b, where

ABLE 26.1 ample Data or Nested o-Factor tudy—aining School Example (class learning scores, coded).

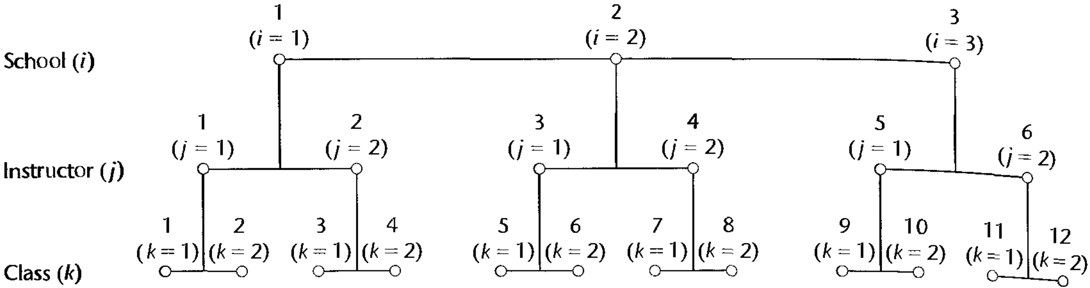
Factor A (school) i	Factor B (instructor) j: 1	2	Average
Atlanta	25	14	
	29	11	
Average	$\bar{Y}_{11.} = 27$	$\bar{Y}_{12.} = 12.5$	$\bar{Y}_{1..} = 19.75$
Chicago	11	22	
	6	18	
Average	$\bar{Y}_{21.} = 8.5$	$\bar{Y}_{22.} = 20$	$\bar{Y}_{2..} = 14.25$
San Francisco	17	5	
	20	2	
Average	$\bar{Y}_{31.} = 18.5$	$\bar{Y}_{32.} = 3.5$	$\bar{Y}_{3..} = 11.00$
		Average	$\bar{Y}_{...} = 15$

FIGURE 26.1 Illustration of Crossed and Nested Factors—Training School Example.



the crossed-out cells represent treatments not studied. Figure 26.2 contains an alternative graphic representation of the nested design for the training school example, including the two replications of the study.

FIGURE 26.2 Graphic Representation of Two-Factor Nested Design—Training School Example.



Nested Design Model

Let Y_{ijk} denote the response for the kth trial when factor A is at the ith level and factor B is at the jth level. We assume that there are n replications for each factor level combination, i.e., $k = 1, \ldots, n$, and that $i = 1, \ldots, a$ and $j = 1, \ldots, b$. Such a study is said to be *balanced* because the same number of factor B levels is nested within each factor A level and the number of replications is the same throughout.

When both factors A and B have fixed effects, an appropriate nested design model is:

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk} \tag{26.7}$$

where:

$\mu_{..}$ is a constant

α_i are constants subject to the restriction $\sum \alpha_i = 0$

$\beta_{j(i)}$ are constants subject to the restrictions $\sum_j \beta_{j(i)} = 0$ for all i

ε_{ijk} are independent $N(0, \sigma^2)$

$i = 1, \ldots, a;\ j = 1, \ldots, b;\ k = 1, \ldots, n$

The expected value and variance of observation Y_{ijk} for nested design model (26.7) with fixed factor effects are:

$$E\{Y_{ijk}\} = \mu_{..} + \alpha_i + \beta_{j(i)} \tag{26.8a}$$

$$\sigma^2\{Y_{ijk}\} = \sigma^2 \tag{26.8b}$$

Thus, all observations have a constant variance. Further, the observations Y_{ijk} are independent and normally distributed for this model.

Random Factor Effects

If both factors A and B have random factor levels, nested design model (26.7) is modified with α_i, $\beta_{j(i)}$, and ε_{ijk} being independent normal random variables with expectations 0 and variances σ_α^2, σ_β^2, and σ^2, respectively. Thus, it is assumed that all $\beta_{j(i)}$ have the same variance σ_β^2. The assumption that all $\beta_{j(i)}$ have the same variance also is made if only factor B is random. It is important to check whether this assumption is appropriate, since it may well be that the mean responses $\mu_{i1}, \mu_{i2}, \ldots$, in one factor A level (plant, school, city, etc.) differ in variability from those in other factor A levels (other plants, schools, cities, etc.). Tests for equality of variances are discussed in Section 18.2.

26.3 Analysis of Variance for Two-Factor Nested Designs

Fitting of Model

The least squares and maximum likelihood estimators of the parameters in nested design model (26.7) are obtained in the usual fashion. Employing our customary notation for sample data in factorial studies, the estimators are:

Parameter	Estimator	
$\mu_{..}$	$\hat{\mu}_{..} = \bar{Y}_{...}$	(26.9a)
α_i	$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$	(26.9b)
$\beta_{j(i)}$	$\hat{\beta}_{j(i)} = \bar{Y}_{ij.} - \bar{Y}_{i..}$	(26.9c)

The fitted values therefore are:

$$\hat{Y}_{ijk} = \bar{Y}_{...} + (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{ij.} - \bar{Y}_{i..}) = \bar{Y}_{ij.} \tag{26.10}$$

and the residuals are:

$$e_{ijk} = Y_{ijk} - \hat{Y}_{ijk} = Y_{ijk} - \bar{Y}_{ij.} \tag{26.11}$$

Sums of Squares

The analysis of variance for nested design model (26.7) is obtained by decomposing the total deviation $Y_{ijk} - \bar{Y}_{...}$ as follows:

$$\underbrace{Y_{ijk} - \bar{Y}_{...}}_{\text{Total deviation}} = \underbrace{\bar{Y}_{i..} - \bar{Y}_{...}}_{A \text{ main effect}} + \underbrace{\bar{Y}_{ij.} - \bar{Y}_{i..}}_{\substack{\text{Specific } B \\ \text{effect when } A \\ \text{at } i\text{th level}}} + \underbrace{Y_{ijk} - \bar{Y}_{ij.}}_{\text{Residual}} \tag{26.12}$$

When we square (26.12) and sum over all cases, all cross-product terms drop out and we obtain:

$$SSTO = SSA + SSB(A) + SSE \tag{26.13}$$

where:

$$SSTO = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{...})^2 \tag{26.13a}$$

$$SSA = bn \sum_i (\bar{Y}_{i..} - \bar{Y}_{...})^2 \tag{26.13b}$$

$$SSB(A) = n \sum_i \sum_j (\bar{Y}_{ij.} - \bar{Y}_{i..})^2 \tag{26.13c}$$

$$SSE = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij.})^2 = \sum_i \sum_j \sum_k e_{ijk}^2 \tag{26.13d}$$

TABLE 26.3 ANOVA Table for Nested Balanced Two-Factor Fixed Effects Model (26.7) (B nested within A).

Source of Variation	SS	df	MS	E{MS}
Factor A	$SSA = bn\sum(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot\cdot})^2$	$a-1$	MSA	$\sigma^2 + bn\frac{\sum\alpha_i^2}{a-1}$
Factor B (within A)	$SSB(A) = n\sum\sum(\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot})^2$	$a(b-1)$	$MSB(A)$	$\sigma^2 + n\frac{\sum\sum\beta_{j(i)}^2}{a(b-1)}$
Error	$SSE = \sum\sum\sum(Y_{ijk} - \bar{Y}_{ij\cdot})^2$	$ab(n-1)$	MSE	σ^2
Total	$SSTO = \sum\sum\sum(Y_{ijk} - \bar{Y}_{\cdot\cdot\cdot})^2$	$abn-1$		

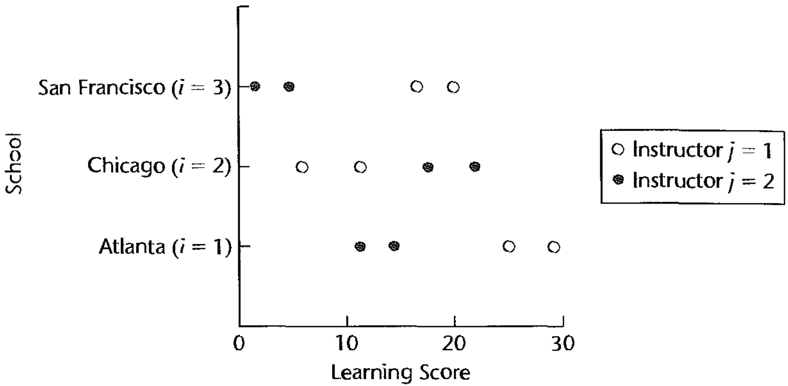


FIGURE 26.3 Dot Plots of Class Learning Scores—Training School Example.

ABLE 26.4 ANOVA for o-Factor ested esign—aining chool xample.

(a) ANOVA Table

Source of Variation	*SS*	*df*	*MS*
Schools (A)	$SSA = 156.5$	2	78.25
Instructors, within schools [$B(A)$]	$SSB(A) = 567.5$	3	189.17
Error (E)	$SSE = 42.0$	6	7.00
Total	$SSTO = 766.0$	11	

(b) Decomposition of SSB(A)

Source of Variation	$SSB(A_i)$	*df*	$MSB(A_i)$
Instructors, Atlanta	210.25	1	210.25
Instructors, Chicago	132.25	1	132.25
Instructors, San Francisco	225.00	1	225.00
Total	567.5	3	

TABLE 26.5 **Expected Mean Squares for Nested Balanced Two-Factor Designs with Random Factor Effects (*B* nested within *A*).**

Mean Square	Expected Mean Square: *A* Fixed, *B* Random	Expected Mean Square: *A* Random, *B* Random
MSA	$\sigma^2 + bn\frac{\sum \alpha_i^2}{a-1} + n\sigma_\beta^2$	$\sigma^2 + bn\sigma_\alpha^2 + n\sigma_\beta^2$
MSB(*A*)	$\sigma^2 + n\sigma_\beta^2$	$\sigma^2 + n\sigma_\beta^2$
MSE	σ^2	σ^2

Test for	Appropriate Test Statistic: *A* Fixed, *B* Random	Appropriate Test Statistic: *A* Random, *B* Random
Factor *A*	*MSA*/*MSB*(*A*)	*MSA*/*MSB*(*A*)
Factor *B*(*A*)	*MSB*(*A*)/*MSE*	*MSB*(*A*)/*MSE*

Estimation of Factor Level Means $\mu_{i.}$

When factor A (fixed effects factor) has significant main effects, there is frequent interest in estimating the factor level means $\mu_{i..}$. The estimated factor level mean $\bar{Y}_{i..}$ is an unbiased estimator of $\mu_{i..}$. As usual for a fixed effects factor, the estimated variance of $\bar{Y}_{i..}$ is based on the mean square in the denominator of the statistic used for testing for factor A main effects, and on the number of cases on which $\bar{Y}_{i..}$ is based. Confidence limits for $\mu_{i.}$ are of the customary form:

$$\bar{Y}_{i..} \pm t(1-\alpha/2; df)s\{\bar{Y}_{i..}\} \tag{26.21}$$

where:

$$s^2\{\bar{Y}_{i..}\} = \frac{MSE}{bn} \qquad df = ab(n-1) \qquad A \text{ and } B \text{ fixed} \tag{26.21a}$$

$$s^2\{\bar{Y}_{i..}\} = \frac{MSB(A)}{bn} \qquad df = a(b-1) \qquad A \text{ fixed, } B \text{ random} \tag{26.21b}$$

Estimation of Treatment Means μ_{ij}

Confidence limits for μ_{ij} are set up in the usual fashion using the t distribution when both factors A and B have fixed effects:

$$\bar{Y}_{ij\cdot} \pm t[1-\alpha/2; (n-1)ab]s\{\bar{Y}_{ij\cdot}\} \tag{26.23}$$

where:

$$s^2\{\bar{Y}_{ij\cdot}\} = \frac{MSE}{n} \tag{26.23a}$$

To make a comparison within any factor A level, we estimate the contrast $L = \sum c_j \mu_{ij}$, where $\sum c_j = 0$, with the estimator $\hat{L} = \sum c_j \bar{Y}_{ij\cdot}$ and employ the confidence limits:

$$\hat{L} \pm t[1-\alpha/2; (n-1)ab]s\{\hat{L}\} \tag{26.24}$$

where:

$$s^2\{\hat{L}\} = \frac{MSE}{n} \sum c_j^2 \tag{26.24a}$$

stimation of Overall Mean $\mu_{..}$

Sometimes there is interest in estimating the overall mean $\mu_{..}$. For the training school example, $\mu_{..}$ is the overall mean learning score for all training schools and all instructors in these schools. The point estimator is $\bar{Y}_{...}$. The confidence limits are constructed utilizing the t distribution as follows:

$$\bar{Y}_{...} \pm t(1-\alpha/2; df)s\{\bar{Y}_{...}\} \tag{26.25}$$

where:

$$s^2\{\bar{Y}_{...}\} = \frac{MSE}{abn} \qquad df = ab(n-1) \qquad A \text{ and } B \text{ fixed} \tag{26.25a}$$

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}, \quad i=1,\dots,a,\ j=1,\dots,b,\ k=1,\dots,n$$

$$\beta_{j(i)} \sim N(0, \sigma_\beta^2) \perp \varepsilon_{ijk} \sim N(0, \sigma^2)$$

Fix: α_i

random: $\beta_{(i)}$

Nested Design

$$E[SSB(A)] = E\left[n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot})^2\right]$$

$$= n \cdot \sum_{i=1}^{a} E\left[\sum_{j=1}^{b} (\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot})^2\right]$$

$$= n \cdot \sum_{i=1}^{a} \left[(b-1) \cdot Var(\bar{Y}_{ij\cdot})\right]$$

$$= n \cdot a \cdot (b-1) \cdot \left(\sigma_\beta^2 + \frac{1}{n}\sigma^2\right)$$

$$= a \cdot (b-1) \cdot (n\sigma_\beta^2 + \sigma^2)$$

$$\frac{1}{b}\sum_{j=1}^{b} \bar{Y}_{ij\cdot} = \bar{Y}_{i\cdot\cdot}$$

$$Var(\bar{Y}_{ij\cdot}) = Var(\bar{Y}_{ij\cdot} - \mu - \alpha_i)$$

$$= Var(\beta_{j(i)} + \bar{\varepsilon}_{ij\cdot})$$

$$= Var(\beta_{j(i)}) + \frac{1}{n} Var(\varepsilon_{ijk})$$

$$= \sigma_\beta^2 + \frac{1}{n}\sigma^2$$

$$E[MSB(A)] = E\left[\frac{SSB(A)}{a \cdot (b-1)}\right] = n\sigma_\beta^2 + \sigma^2$$

$$Var(\bar{Y}_{i\cdot\cdot}) = Var(\bar{Y}_{i\cdot\cdot} - E\bar{Y}_{i\cdot\cdot})$$

$$= Var\left(\frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n}(\beta_{j(i)} + \varepsilon_{ijk})\right)$$

$$= Var\left(\frac{1}{b}\sum_{j=1}^{b}\beta_{j(i)} + \frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n}\varepsilon_{ijk}\right)$$

$$= Var\left(\frac{1}{b}\sum_{j=1}^{b}\beta_{j(i)}\right) + Var\left(\frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n}\varepsilon_{ijk}\right)$$

$$= \frac{1}{b}\sigma_\beta^2 + \frac{1}{bn}\sigma^2 = \frac{1}{bn}(n\sigma_\beta^2 + \sigma^2) = \frac{E[MSB(A)]}{bn}$$

$$\therefore\ S^2(\bar{Y}_{i\cdot\cdot}) = \frac{MSB(A)}{bn}$$

the square of standard error for A fixed, B random

and C.I. for $\mu_i = \mu + \alpha_i$ is:

$$\bar{Y}_{i\cdot\cdot} \pm t_{\frac{\alpha}{2}}(a \cdot (b-1)) \cdot S(\bar{Y}_{i\cdot\cdot})$$