CHAPTER 7

Nested Designs

Nested designs and the analysis of data from experiments that use these designs are considered to be complicated. Indeed, Reece (2003, p. 384) stated the following regarding nested designs and their analysis using Design-Expert, Version 6.0 (DX6): "The analyses of these designs are complicated at best, but are possible with DX6."

The problem is compounded somewhat by the fact that most books on experimental design devote comparatively little space to the subject, and hardly any of the leading design books have a chapter devoted exclusively to nested designs. Exceptions include Mason, Gunst, and Hess (2003) and Yandell (1997), each of which has two chapters on nested designs. Many design books barely even mention nested designs, so it can be somewhat difficult to gain an in-depth knowledge of the subject without consulting a small number of good sources of information.

One example of the confusion that exists regarding nested designs is the following statement, which can be found at www.sas.org/E-Bulletin/2002-03-08/features/body.html : "Nested designs are not an exclusive category, because a full 2^2 factorial design is also a nested design." As was shown in Section 4.1, in a 2^2 factorial design each level of each factor occurs with each level of the other factor in the design. This is not the case when a nested design is used. That is, the levels of the factors are not crossed. Therefore, the statement is in error.

Nested designs are also called hierarchical designs, with the term emanating from the hierarchical (i.e., "nesting") structure between the factors. There can be either strict nesting of factors or a combination of nesting and a factorial arrangement. When the latter exists, the design is called a *nested factorial design*.

Assume that we have a manufacturing experiment in which temperature at some stage of a process is varied in an experiment to see if it has a significant effect on process yield. Three temperatures are used and each is used for two consecutive weeks, so the experiment runs for six weeks. The factor of interest is temperature, but there might be a week effect (we hope not), which would be a nuisance factor.

Assume that we are interested in looking at both weeks and temperature. What type of design is this? In a factorial design each level of every factor is "crossed" with each level of every other factor. Thus, if there are a levels of one factor and b levels of a second factor, there are then ab combinations of factor levels. Since there are actually six weeks involved in the temperature-setting experiment, there would have to be $6 \times 3 = 18$ combinations for it to be a cross-classified design. Since there are obviously only six combinations of weeks and temperatures, it clearly cannot be a cross-classified (factorial) design.

Then what type of design is this? It is actually a *nested factor design* as weeks are "nested" within temperature. The corresponding model is

$$Y_{ijk} = \mu + A_i + B_{j(i)} + \varepsilon_{k(ij)} \qquad i = 1, 2, 3 \quad j = 1, 2 \quad k = 1, 2, 3, 4, 5 \qquad (7.1)$$

where i designates the temperature, j the week, and k the replicate factor (days in this case). Further, $j(i)$ indicates that weeks are nested within temperature, and $k(ij)$ indicates that the replicate factor is nested within each (i, j) combination.

Notice that the model does not contain an interaction term. There can be no interaction terms in the model when a strictly nested design is used since an interaction cannot be computed unless all possible combinations of the levels of factors that comprise the interaction are used. Even though interactions cannot be computed, this doesn't mean that they don't exist. The users of these designs simply must assume that they don't exist, but assuming that interactions don't exist certainly doesn't cause them to not exist. It should be noted that it is possible to have a design with a combination of nesting and a factorial structure, as is explained in Section 7.4.

In some applications there can, by definition, be no interaction. For example, assume we have two different types of machines and three types of heads that would fit on one machine but not on the other one. There can be no interaction between heads and machines because the heads cannot be swapped between machines. So for this type of application the assumption of no interaction is certainly true. (There is more discussion of the machines–heads scenario in later sections and in Exercise 7.12.)

On the other hand, assume that a drug-testing experiment is conducted using two hospitals and four drugs. Assume that two drugs are used at one hospital and the other two drugs used at the other hospital. Here we do have to assume no interaction because the drugs should be interchangeable between the hospitals. Assume that the variation in response values would have been much less (for whatever reasons) if the first two drugs had been used at the second hospital instead of the first hospital. Then the assumption of no interaction would not be valid and the results of the experiment could be misleading.

Nested designs are used in a variety of applications, including the chemical industry (see, e.g., Lamar and Zirk, 1991 who gave illustrative case studies of a three-level nested design and a four-level nested design) and manufacturing (see, e.g., Liu and Batson, 2003). More specifically, the use of these designs in process control and process variation studies has been discussed by Bainbridge (1965), Sinibaldi (1983), Snee (1983), and Pignatiello (1984). The designs are also used in agricultural and psychological research, in addition to many applications in biology. Vander Heyden, De Braekeleer, Zhu, Roets, Hoogmartens, De Beer, and Massart (1999) described the use of nested designs in ruggedness testing in pharmaceutical work and proposed an alternative analysis method for use in ruggedness testing.

One of the primary objectives in the use of nested designs is to remove uncontrolled variation due to a priori differences in primary sampling units from the experimental error. Another reason is that in biological experiments, in particular, there is often a desire to be able to make inferences regarding hierarchically arranged environments, habitats, or species.

Often there is no choice between a nested design and a crossed design because circumstances dictate use of the former. (For example, a head might be usable on only one particular machine, the Florida Everglades can obviously be found only in Florida and not in any other state, etc.)

As is discussed in Chapter 9, it is important to recognize a split-plot structure when it exists, since analyzing the data as if this structure does not exist can cause misleading results. Similarly, it is important to distinguish a nested design structure from other structures. We consider this issue with the following example.

Example 7.1

Pignatiello (1984) discussed a scenario in which an engineer was working for a military contractor and sampling from lots was performed. Specifically, five parts were randomly sampled from each of two lots, and two measurements were made on each part since measurement error was known to exist. The engineer apparently decided to use the average measurement on each part if there was no difference in the variances between the two lots. In other words, measurement error would be removed from the analysis. But equality of variances, if it exists, does not obviate an analysis that uses the measurement error and the fact that there are repeat readings on each part. Furthermore, this raises the question of whether or not this is really a repeated measures design (Chapter 11) and whether the data should be analyzed as such. With a repeated measures design each experimental unit receives more than one treatment. That doesn't happen here, as measuring a part twice does not mean that two treatments are being applied. Therefore, this is not a repeated measures design. There might seem to be a type of nesting involved here, however, since measurement error on each part is nested within that part, but that is hardly different from saying that the variability within each factor level in a one-factor design is nested within each level.

We may note, however, that what the engineer wanted to do is an example of a practitioner trying to collapse data for simplicity, which is generally not a good idea. There is information in numbers, just as there is more information in a cholesterol reading of 217 than simply reporting "over 200." There were no data reported in Pignatiello (1984) since the author simply reported an inquiry, and so no analysis can be given here.

As stated above, each experimental unit receives more than one treatment when a repeated measures design is used, whereas each experimental unit receives only one treatment when a nested design is used.

Example 7.2

An example of a purported nested design appears on the Internet (http://www. psych-stat.missouristate.edu/introbook/sbk23m.htm). The experiment consisted of six males and nine females and a t-test was performed to test for the equality of means, with the response variable being "finger-tapping speed" (no other details were given). The statement is made that "The design is necessarily nested because each subject has only one score and appears in a single treatment condition." There is no treatment, however, at least not in the usual sense. To state that this is an example of a nested design would be similar to saying that one-factor ANOVA is a nested design because each experimental unit receives only one treatment. Furthermore, this scenario is really more of an observational study than a designed experiment. Obviously there is no randomization involved because if we think of sex as the factor and the person as the experimental unit, we obviously can't assign sex to the person.

7.1 VARIOUS EXAMPLES

There are many good examples of nested designs. An experiment might be conducted with the objective of improving process yield, with the experiment involving five machines and five operators. Unless each operator uses each machine during the experiment (which might be both impractical and cost prohibitive) the operator effect will be nested under the machine effect.

As another example, consider an experiment to study the strain readings of glass cathode using four different machines. Assume that there are different heads that will fit on each machine, but will not fit on any of the other three machines. One head is to be randomly selected for each machine. Since each head that is selected cannot be used on any of the other machines, heads are thus nested within machines, so the head factor is a nested factor.

Another example, which has been found on the Internet, is that of a forestry experiment that involves three forests, five trees that are sampled within each forest, and five seedlings that are sampled for each tree. Thus, the tree effect, if it exists, is nested within forests and if a seedling effect exists, it is nested within trees. The ANOVA table setup would be as follows.

Source	df
Forests	2
Trees within forests	12
Seedlings within trees	60
TOTAL	74

Hypothesis testing is performed by constructing F-tests using the rules that apply to factorial designs regarding fixed and random factors. That is, Forests are tested against Seedlings if both Forests and Trees are fixed; tested against Trees if the latter is random and Forests are fixed, or both factors are random; and Trees are of course tested against Seedlings regardless of whether the two factors are fixed or random.

7.2 SOFTWARE SHORTCOMINGS

Whereas data from experiments with nested designs can be analyzed using SAS Software, trying to do the analysis with most other software packages is either difficult or impossible. For example, Reece (2003) points out, in referring to Design-Expert: "By manipulating it's analysis capabilities, a user can analyze nested designs, but it does not support generating them directly." (This comment was in regard to Version 6.0.5.) Similarly, there is no mention of nested designs in D. o. E. Fusion Pro. Data may be analyzed as having come from a nested design in JMP, however, by using the FitModel dialog with the "Nest" button. Users of nested designs may wish to consider RS/Discover, as Reece (2003, p. 364) stated, "RS/Discover is one of the few products tested that provides support for nested designs involving a number of random effects."

Moderately extensive nested design analysis capability is available in MINITAB, but the key word here, as it is in JMP, is *analysis*. That is, data from the use of a nested design can be analyzed, both for a fully nested design and for an unbalanced nested design. *However*, the "fully nested ANOVA" routine in MINITAB is only for random factors and thus cannot be used if there is at least one fixed factor. When the latter is true, the general linear model (GLM) routine must be used, and the model and the nesting structure must be specified. That is, if the two factors used in an experiment are Machines and Operators and the latter is nested under the former, the terms in the model would have to be specified as "Machines Operators (Machines)," regardless of whether command mode or menu mode is used. Since it is not uncommon for one of the factors in a nested experiment to be fixed (such as the first factor, under which the other factors are nested), many MINITAB users will have to use the GLM capability for analysis.

Data from a gauge R & R (reproducibility and repeatability) study can also be analyzed with nested factors, as will often occur, such as operators being nested under machines rather than being crossed with machines. (It isn't necessary, in MINITAB, to specify the nesting structure for this routine as that is implied by the designation of the appropriate columns for "Operators" and "Part numbers.")

The situation regarding software isn't quite as bleak as it may seem, however, as is explained in the next section.

7.2.1 A Workaround

Experimenters and others who are using something other than JMP, SAS, or MINITAB can still use their software, although some of the work will have to be performed manually. Because of this, we need to look at some sum of squares results in certain detail.

Accordingly, assume that there are two factors and factor B is nested under factor A. The latter has two levels and factor B is at two levels for each level of factor A.

Since the levels are different, this means that factor B has four levels. Thus there are four combinations of levels of A and B. With a factorial arrangement there would have been eight combinations so this obviously isn't a factorial arrangement. Assume further that there is only one observation per treatment combination, so there are four observations altogether.

We can use the computations that would be used *if* this were a 2^2 design to arrive at the sums of squares for the nested design. That is, we will pretend that the third and fourth levels of factor B are the same as the first and second levels. Recall the formulas for estimating the B effect and the $A \times B$ interaction effect that were given in Section 4.3, and also recall from Section 4.5 that for an unreplicated 2^2 design the sum of squares for a particular effect is equal to the square of the effect estimate.

Squaring the effect estimates for B and the $A \times B$ interaction and adding them together produces $[(A_1B_1)^2 + (A_1B_2)^2 + (A_2B_1)^2 + (A_2B_2)^2]/2 - A_1B_1A_1B_2 - A_2B_1A_2B_2$.

Now let's consider how we would compute the sum of squares for a nested factor, and specifically B nested within A. We are interested in the variability of the response using the two levels of B at each level of A. Therefore, we want to compute $\sum_{i=1}^{2}\sum_{j=1}^{2}(A_iB_j - \overline{A}_i)^2$, with $\overline{A}_i$ denoting the average response at the ith level of A. Expanding and simplifying this expression produces the same expression given above for the sum of B and $A \times B$, as the reader is asked to show in Exercise 7.2. This means that software that will handle a 2^2 design (as virtually all general purpose statistical software packages will) can be used to produce the sum of square components that are needed for the analysis of the two-factor nested design. Of course it isn't particularly satisfying to have to use software to produce components that then must be combined manually to provide the ANOVA table, but this approach will have to be used by many practitioners, depending upon the software to which they have access.

We will illustrate the necessary computations with the following example.

Example 7.3

Assume that the two major hospitals in a small city are selected for an experiment involving four drugs, with these drugs being the only ones of interest. For the sake of illustration we will assume that the 12 patients who will be involved in the experiment (3 for each hospital–drug combination) are essentially homogenous in regard to any physical characteristics that could affect the response values. The layout and the (coded) response values are given below, with H_i denoting the ith hospital, $i = 1, 2$; and D_i denotes the ith drug, $i = 1, 2, 3, 4$.

H_1		H_2	
D_1	D_2	D_3	D_4
(5, 7, 4)	(6, 8, 5)	(8, 9, 6)	(9, 8, 10)

If we analyze the data as having come from an experiment with a replicated 2^2 design, we obtain the sum of squares given in the following ANOVA table.

Source	DF	SS
H	1	18.750
D	1	4.083
H*D	1	0.083
Error	8	16.000
Total	11	38.917

It follows that the ANOVA table for the nested design is

Source	DF	SS
H	1	18.750
D(H)	2	4.166
Error	8	16.000
Total	11	38.917

Notice that the degrees of freedom for $D(H)$ after the pooling is as it should be because D has one degree for each level of H since there are two levels of D at each level of H. If we calculated the sum of squares for $D(H)$ directly, we would have done so by computing $3\sum_{i=1}^{2}\sum_{j=1}^{2}(\overline{H_i D_j} - \overline{H}_i)^2$, with the "3" resulting from the fact that there are three observations for each hospital–drug combination. The reader is asked to show in Exercise 7.3 that this produces the sum of squares for $D(H)$ shown in the ANOVA table.

Source	DF	SS	MS	F	p
H	1	18.750	18.750	9.375	0.016
D(H)	2	4.166	2.083	1.04	0.396
Error	8	16.000	2.000		
Total	11	38.917			

The F-statistics are computed using the fact that both factor H and factor D are fixed. We see that the hospital effect is significant at the .01 level, but the drug effect is not significant, which is the reverse of what an experimenter would want to see for this experiment since the conclusion is that the drugs do not differ in effectiveness.

Of course it is bothersome to have to obtain the F-statistics and p-values manually. We can avoid the manual construction of F-tests by using JMP or SAS. The output from JMP for this example is as follows.

```
Summary of Fit

RSquare        0.588865
RSquare Adj        0.43469
Root Mean Square Error      1.414214
Mean of Response      7.083333
Observations (or Sum Wgts) 12
```

```
                     Analysis of Variance

 Source    DF   Sum of Squares   Mean Square   F Ratio   Prob > F
 Model      3   22.916667          7.63889     3.8194    0.0575
 Error      8   16.000000          2.00000
 C.Total   11   38.916667

      Effect Tests

Source              Nparm  DF   Sum of Squares  F Ratio   Prob > F
Hospital                1   1      18.750000     9.3750    0.0155
Drugs [Hospital]        2   2       4.166667     1.0417    0.3962
```

The output using MINITAB is as follows.

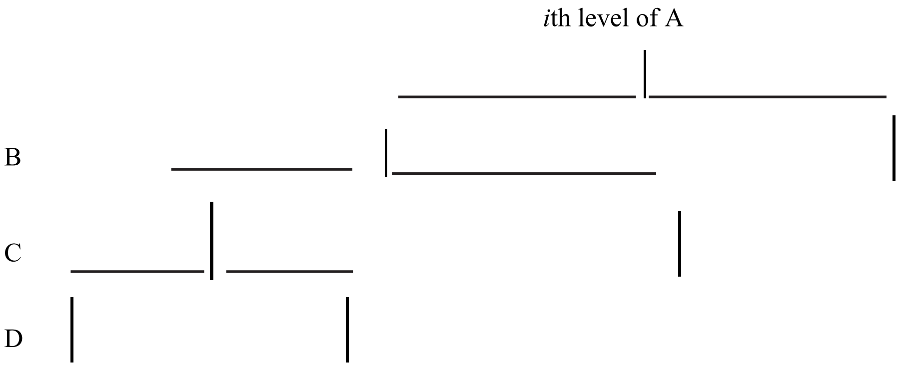
```
    General Linear Model: Response versus Hospital, Drug

    Factor           Type    Levels  Values
    Hospital         fixed        2  1, 2
    Drug(Hospital)   fixed        4  1, 2, 3, 4

Analysis of Variance for Response, using Sequential SS for Tests

Source          DF   Seq SS   Adj SS   Seq MS      F       P
Hospital         1   18.750   18.750   18.750   9.38   0.016
Drug(Hospital)   2    4.167    4.167    2.083   1.04   0.396
Error            8   16.000   16.000    2.000
Total           11   38.917

S = 1.41421 R-Sq = 58.89% R-Sq(adj) = 43.47%

Expected Mean Squares, using Sequential SS

                    Expected Mean
                    Square for
   Source           Each Term
1  Hospital         (3) + Q[1, 2]
2  Drug(Hospital)   (3) + Q[2]
3  Error            (3)

Error Terms for Tests, using Sequential SS

                                        Synthesis
                                        of Error
   Source           Error DF  Error MS  MS
1  Hospital             8.00     2.000  (3)
2  Drug(Hospital)       8.00     2.000  (3)
```

7.3 STAGGERED NESTED DESIGNS

Even though a nested design does not have a factorial structure, the degrees of freedom does have a “product structure” that can result in a large number of degrees of freedom—more than is necessary—for effects at the bottom of the hierarchical structure. This is analogous to what happens with interactions in factorial designs with more than two levels, as for example, the *ABC* interaction in a 4^3 design has 27 degrees of freedom, far more than is necessary to estimate an effect that probably isn't real. Although there are of course no interactions in nested designs, the degrees of freedom for effects in nested designs has a somewhat similar product structure. Specifically, the degrees of freedom for a given effect at least two levels from the top is the product of the number of levels of the factors above it in the hierarchical structure multiplied times one less than the number of levels of that effect. For example, if factor C is nested under factor B, which in turn is nested under factor A and all factors have four levels, factor C has $(4)(4)(4-1) = 48$ degrees of freedom and the error degrees of freedom would be much larger than this, resulting in a sample size that may be cost prohibitive. Thus, nested designs can be inefficient, as discussed, for example, by Khuri and Sahai (1985).

The largest reductions in this number would occur when the levels of factor A and/or factor B are reduced, but this might not be acceptable if an experimenter has a certain number of levels in mind to be investigated.

Various alternatives to a nested design have been discussed in the literature, some of which are covered by Bainbridge (1965). A *staggered nested design* is one practical alternative, which is recommended by Bainbridge (1965). With this type of design, there is not full nesting. For example, consider the following staggered design layout given by Smith and Beverly (1981).



Factors A, B, C, and D might represent raw material lots, batches, samples, and measurements, respectively. Notice that two samples are obtained from only one of the two batches (with one sample taken from the other batch), with one of the two samples having a repeat measurement. Thus, there are only four observations for the ith level of A with this staggered nested design, whereas there would be eight observations with a regular nested design.

Assume that there are three levels of A. The degrees of freedom breakdown is then

Source	df
A	2
B (within A)	3
C (within B)	3
D (error)	3
Total	11

Since a staggered nested design is an unbalanced design, it can be analyzed by software that handles unbalanced designs, such as the GLM capability in MINITAB. Although these designs are rarely used in biology, Cole (2001) gave an example of their use in such an application.

7.4 NESTED AND STAGGERED NESTED DESIGNS WITH FACTORIAL STRUCTURE

It is possible to construct designs that have a factorial structure for some factors and a nested structure for other factors, as stated previously. Similarly, it is also possible to construct a design that has a staggered nested structure for some factors and a factorial structure for other factors. Smith and Beverly (1981) discuss these designs and the models to which the designs correspond. Khattree and Naik (1995) gave statistical tests for random effects for these designs; Khattree, Naik, and Mason (1997) discussed the estimation of variance components for the designs; and Naik and Khattree (1998) gave a computer program for estimating the variance components. Ojima (1998) gave certain theoretical results.

As in Eq. (7.1), the way in which the model is written gives the nesting structure. For example, the model

$$Y_{ijkl} = \mu + A_i + B_j + C_{k(ij)} + \varepsilon_{l(ijk)} \tag{7.2}$$

shows that the effect of factor C is nested within the AB combinations and that the error term also has a nested structure, as in Eq. (7.1). Thus, the model in Eq. (7.2) has both a factorial structure and a nested structure. Although it was stated at the beginning of the chapter that the model for a nested design cannot contain any interaction terms, this is assuming that all of the factors are nested. It is possible to have interaction terms involving factors that have factorial structure, so an AB_{ij} term in Eq. (7.2) would be permissible since factors A and B have factorial structure.

7.5 ESTIMATING VARIANCE COMPONENTS

To this point in the chapter, only inference using ANOVA has been presented. Random factors generally occur when nested designs are used, as factors that are nested are generally random, and sometimes all factors are random, as in Exercise 7.9. (When all factors are random, the NESTED command in MINITAB can be used for the analysis.)

Since variance reduction is important in quality improvement work, for example, having variance component estimates for factors in an experiment can be quite useful. Estimates will frequently be negative, however, which is a problem since variances are of course never negative.

The following output will be used to illustrate how this can occur.

```
        Analysis of Variance for Y

Source   DF        SS           MS          F        P
A         2   30236.7222   15118.3611   166.696   0.001
B         3     272.0833      90.6944     0.694   0.573
C        12    1569.0000     130.7500     7.679   0.000
D        18     306.5000      17.0278
Total    35   32384.3056

           Variance Components

                        % of
  Source   Var Comp.   Total    StDev
    A      1252.306    94.43   35.388
    B        -6.676*    0.00    0.000
    C        56.861     4.29    7.541
    D        17.028     1.28    4.126
   Total   1326.194            36.417

  * Value is negative, and is estimated by zero.

            Expected Mean Squares

   1   A   1.00(4) + 2.00(3) + 6.00(2) + 12.00(1)
   2   B   1.00(4) + 2.00(3) + 6.00(2)
   3   C   1.00(4) + 2.00(3)
   4   D   1.00(4)
```

The last part of the output is used in combination with the first section to obtain the variance component estimates that are shown in the middle section. To illustrate, $E(\mathrm{MS}_D) = \sigma_D^2$ so $\widehat{\sigma}_D^2 = \mathrm{MS}_D = 17.028$. $E(\mathrm{MS}_C) = \sigma_D^2 + 2\sigma_C^2$. Using this last equation as the basis for solving for $\widehat{\sigma}_C^2$ produces $\widehat{\sigma}_C^2 = (130.75 - 17.028)/2 = 56.861$. We encounter a problem when we try to solve for $\widehat{\sigma}_B^2$, however. $E(\mathrm{MS}_B) = \sigma_D^2 + 2\sigma_C^2 + 6\sigma_B^2$, as shown. Using the estimates for σ_D^2 and σ_C^2, we obtain $\widehat{\sigma}_B^2 = (90.6944 - 17.028 - 2(56.861)/6 = -6.676$. Notice that the output sets the estimate to zero. The negative estimate is caused by the large value of $\widehat{\sigma}_C^2$, especially relative to the value of MS_B. The large value should be investigated and in fact was investigated, as is explained in the particular chapter exercise for which the analysis of the data results in this computer output.

Readers interested in hand computation of variance component estimates are referred to Nelson (1995a), with confidence intervals for the corresponding standard deviations given in Nelson (1995b). Variance component estimation for up to three-level nesting is covered by Searle, Casella, and McCulloch (1992), and for confidence interval construction of variance components readers are referred to Burdick and Graybill (1992). For applications of variance components estimation with nested designs in the semiconductor industry, see Jensen (2002); see also Ankenman, Liu, Karr, and Picka (2002). Another useful reference, but now out of print, is Rao and Kleffe (1988), and Vardeman and Wendelberger (2005) provide some instruction in the estimation of variance components for an unbalanced two-factor nested design.

7.6 ANOM FOR NESTED DESIGNS?

Analysis of Means (ANOM) was used initially in Section 2.2 and subsequently used to analyze data from various types of designs. As was stated in Section 3.1.5 and also stated by Nelson (1993), ANOM can be used with any complete design. The factors must be fixed, however, and with a nested design this requirement will hardly ever be met.

Gonzalez-de la Parra and Rodriguez-Loaiza (2003) proposed a ANOM procedure for nested designs with random factors. The proposed method should be viewed as only an ad hoc procedure, however, for the following reasons. The authors simply applied ANOM for a single factor to each of the factors separately and then proceeded as one would do if the factors were fixed. They interpreted the charts by looking for "... patterns (random or systematic) of the subgroup means across the factors...." Thus, this is not a well-defined statistical procedure, nor can it be since ANOM is not applicable to random factors. So even though the use of these charts was beneficial in the application that was described, such an analysis would be best used only as a graphical supplement to accepted procedures.

7.7 SUMMARY

Nested designs have been used in industrial applications for decades, and medical, agricultural, and biological applications are also natural, as are applications in psychological research. Applications in other areas, such as aquatics, have been slow coming, however. Staggered nested designs are a practical alternative to nested designs, as the former provides the opportunity to have a reasonable number of degrees of freedom for factors at the top of the hierarchical structure without having a high number of degrees of freedom at the bottom. Unfortunately, there are some software shortcomings, detailed in Section 7.2, that can hinder the effective use of nested designs and analysis of the data from experiments in which these designs are used. Consequently, experimenters need to be aware of the limitations of specific statistical software.

REFERENCES

Ankenman, B. E., H. Liu, A. F. Karr, and J. D. Picka (2002). A class of experimental designs for estimating a response surface and variance components. *Technometrics*, **44**(1), 45–54.

Bainbridge, T. R. (1965). Staggered nested designs for estimating variance components. *Industrial Quality Control*, **22**, 12–20.

Burdick, R. K. and F. A. Graybill (1992). *Confidence Intervals on Variance Components*. New York: Marcel Dekker.

Cole, R. G. (2001). Staggered nested designs to identify hierarchical scales of variability. *New Zealand Journal of Marine and Freshwater Research*, **35**, 891–896. (This article is available at http://www.rsnz.org/publish/nzjmfr/2001/78.pdf.)

Croarkin, C. and P. Tobias, eds. (2002). *NIST/SEMATECH e-Handbook of Statistical Methods* (http://www.itl.nist.gov/div898/handbook), a joint effort of the National Institute of Standards and Technology and International SEMATECH.

Gonzalez-de la Parra, M. and P. Rodriguez-Loaiza (2003). Application of analysis of means (ANOM) to nested designs for improving the visualization and understanding of the sources of variation of chemical and pharmaceutical processes. *Quality Engineering*, **15**(4), 663–670.

Gruska, G. F. and M. S. Heaphy (2000). Designing experiments—an Overview. *ASQC Statistics Division Newsletter*, **19**(2), 8–13.

Hicks, C. R. (1956). Fundamentals of Analysis of Variance, Part III: Nested designs in analysis of variance. *Industrial Quality Control*, **13**(4), 13–16.

Jensen, C. R. (2002). Variance component calculations: Common methods and misapplications in the semiconductor industry. *Quality Engineering*, **14**, 645–657.

Khattree, R. and D. N. Naik (1995). Statistical tests for random effects in staggered nested designs. *Journal of Applied Statistics*, **22**, 495–505.

Khattree, R., D. N. Naik, and R. L. Mason (1997). Estimation of variance components in staggered nested designs. *Journal of Applied Statistics*, **24**, 395–408.

Khuri, A. I. and H. Sahai (1985). Variance components analysis: A selective literature review. *International Statistical Review*, **53**, 279–300.

Lamar, J. M. and W. E. Zirk (1991). Nested designs in the chemical industry. In *Annual Quality Congress Transactions*, American Society for Quality Control, Milwaukee, WI. pp. 615–622.

Liu, S. and R. G. Batson (2003). A nested experiment design for gauge gain in steel tube manufacturing. *Quality Engineering*, **16**(2), 269–282.

Mason, R. L., R. F. Gunst, and J. L. Hess (2003). *Statistical Design and Analysis of Experiments: With Applications to Engineering and Science*, 2nd ed. New York: Wiley.

Naik, D. N. and R. Khattree (1998). A computer program to estimate variance components in staggered nested designs. *Journal of Quality Technology*, **30**, 292–297.

Nelson, P. R. (1993). Additional uses for the analysis of means and extended tables of critical values. *Technometrics*, **35**(1), 61–71.

Nelson, L. S. (1995a). Using nested designs: I. Estimation of standard deviations. *Journal of Quality Technology*, **27**(2), 169–171.

Nelson, L. S. (1995b). Using nested designs: II. Confidence limits for standard deviations. *Journal of Quality Technology*, **27**(3), 265–277.

Ojima, Y. (1998). General formulae for expectations, variances and covariances of the mean squares for staggered nested designs. *Journal of Applied Statistics*, **25**, 785–799.

Pignatiello, J. (1984). Two-stage nested designs. *ASQC Statistics Division Newsletter*, **6**(1), September. (This can be accessed by ASQ members at http://www.asq.org/forums/statistics/newsletters/vol_6_no_1_sept1984.pdf.)

Rao, C. R. and J. Kleffe (1988). *Estimation of Variance Components and Applications*. Amsterdam: North-Holland.

Reece, J. E. (2003). Software to support manufacturing systems. In *Handbook of Statistics*, Vol. 22, Chapter 9 (R. Khattree and C. R. Rao, eds.). Amsterdam: Elsevier Science B. V.

Salvia, A. A. and W. C. Lasher (1989). The power of the *F*-test in two-factor nested experiments. *Journal of Quality Technology*, **21**(1), 20–23.

Searle, S. R., G. Casella, and C. E. McCulloch (1992). *Variance Components*. New York: Wiley.

Sinibaldi, F. J. (1983). Nested designs in process variation studies. In *ASQC Annual Quality Congress Transactions*, pp. 503–508. American Society for Quality Control, Milwaukee, WI.

Smith, J. R. and J. M. Beverly (1981). The use and analysis of staggered nested factorial designs. *Journal of Quality Technology*, **13**(3), 166–173.

Snee, R. D. (1983). Graphical analysis of process variation studies. *Journal of Quality Technology*, **15**(2), 76–88.

Vander Heyden, Y., K. De Braekeleer, Y. Zhu, E. Roets, J. Hoogmartens, J. De Beer, and D. L. Massart (1999). Nested designs in ruggedness testing. *Journal of Pharmaceutical and Biomedical Analysis*, **20**(6), 875–887.

Vardeman, S. B. and J. Wendelberger (2005). The expected sample variance of uncorrelated random variables with a common mean and some applications in random unbalanced effects models. *Journal of Statistical Education*, **13**(1). (This is an online journal available at http://www.amstat.org/publications/jse.)

Yandell, B. (1997). *Practical Data Analysis for Designed Experiments*. London: Chapman & Hall.

EXERCISES

7.1 Describe an experiment from your field of study for which the use of a nested design would be appropriate.

7.2 Derive the expression for $SS_{B(A)}$ that was given in Section 7.2.1 and similarly derive the expression for $SS_{A\times B} + SS_B$ and show their equality.

7.3 Compute directly the sum of squares for $D(H)$ in Example 7.3.

7.4 Consider the schematic diagram of a nested design given in Vardeman and Wendelberger (2005, references). Determine the number of degrees of freedom for error.

7.5 In listing advantages and disadvantages of nested designs Gruska and Heaphy (2000, references) stated the following: “DISADVANTAGE—Nesting is a specialized design appropriate only when some hierarchical structure is present.” Respond to that statement.

7.6 Salvia and Lasher (1989, references) gave the following example of data from a two-factor nested design. A food processor produces 12 batches of strawberry jam per day. Each batch contains 100 cases of 24 jars. The response variable is the sugar content of the product. Four batches are randomly selected from the 12; three cases are then randomly selected, and then two jars are sampled. The data are given below.

	Batch			
Case	1	2	3	4
1	5.821	5.991	6.123	6.215
	5.815	5.997	6.133	6.203
2	5.772	5.973	6.028	6.143
	5.778	5.947	6.063	6.158
3	5.763	5.924	6.025	6.045
	5.746	5.937	6.018	6.039

Perform an appropriate analysis and state your findings to the head statistician and to the chief operating officer.

7.7 Smith and Beverly (1981, references) gave the following example. A product in pellet form is received in bulk form in hopper-type trailers. An experiment was performed to estimate the between-trailers (long term) variability and the within-trailer (short term) variability of a quality characteristic, impurities. Ten trailers were used. Measurement precision was also to be assessed and the two laboratories used in the experiment were to be compared. Four samples were obtained from each trailer, with two assigned to one laboratory and two assigned to the other. Each laboratory made three tests—two replicate measurements on the same sample and a single measurement on a second sample. Within this restriction, samples were assigned randomly to the laboratories. Thus, there are 60 observations: 10*2*3. The data are given below.

Trailer	Lab	Sample	Measurement	Trailer	Lab	Sample	Measurement
1	1	1	47.06	6	1	1	46.99
1	1	1	44.37	6	1	1	50.87
1	1	2	49.30	6	1	2	51.87
1	2	1	47.40	6	2	1	52.14
1	2	1	47.80	6	2	1	49.56
1	2	2	50.43	6	2	2	48.03
2	1	1	47.43	7	1	1	47.49
2	1	1	50.35	7	1	1	51.55
2	1	2	50.42	7	1	2	58.57
2	2	1	50.43	7	2	1	51.61
2	2	1	53.07	7	2	1	49.86
2	2	2	49.18	7	2	2	46.32
3	1	1	48.90	8	1	1	47.41
3	1	1	48.05	8	1	1	47.63

3	1	2	50.64	8	1	2	48.63
3	2	1	52.52	8	2	1	48.46
3	2	1	50.38	8	2	1	46.14
3	2	2	47.64	8	2	2	47.41
4	1	1	52.32	9	1	1	48.37
4	1	1	52.26	9	1	1	51.03
4	1	2	53.47	9	1	2	50.15
4	2	1	47.39	9	2	1	50.53
4	2	1	50.73	9	2	1	47.82
4	2	2	54.49	9	2	2	49.37
5	1	1	46.53	10	1	1	54.80
5	1	1	45.60	10	1	1	51.57
5	1	2	53.98	10	1	2	54.52
5	2	1	48.07	10	2	1	53.02
5	2	1	47.59	10	2	1	51.95
5	2	2	46.50	10	2	2	50.50

Perform an appropriate analysis and state your conclusions.

7.8 Assume that there are five levels of factor A for the schematic diagram in Section 7.3, instead of three as was assumed in the example. What will be the number of degrees of freedom for error?

7.9 Snee (1983, references) gave an example of a nested design with four factors, with the factors being Operator, Specimen, Run, and Analysis. There were three operators, two specimens that were nested under each level of operator, three levels of run that were nested under each level of specimen, and two levels of analysis that were nested under each level of run. All factors were random. Since the description was sketchy, this may not have been an actual experiment, or perhaps the description was sketchy for proprietary reasons. The data are given below.

Operator	Specimen	Run	Analysis	Response Value
1	1	1	1	156
1	1	1	2	154
1	1	2	1	151
1	1	2	2	154
1	1	3	1	154
1	1	3	2	160
1	2	4	1	148
1	2	4	2	150
1	2	5	1	154
1	2	5	2	157
1	2	6	1	147
1	2	6	2	149
2	3	7	1	125
2	3	7	2	125
2	3	8	1	94

2	3	8	2	95
2	3	9	1	98
2	3	9	2	102
2	4	10	1	118
2	4	10	2	124
2	4	11	1	112
2	4	11	2	117
2	4	12	1	98
2	4	12	2	110
3	5	13	1	184
3	5	13	2	184
3	5	14	1	172
3	5	14	2	186
3	5	15	1	181
3	5	15	2	191
3	6	16	1	172
3	6	16	2	176
3	6	17	1	181
3	6	17	2	184
3	6	18	1	175
3	6	18	2	177

(a) Since all factors are random, interest would be focused on variance component estimation. Snee (1983) makes the point that, in general, atypical observations can inflate variances. Do any of these observations seem unusual?

(b) Perform appropriate analyses and draw conclusions regarding the sources of variation.

(c) You will notice a problem that occurs with a variance component estimate. Is the cause of the problem apparent from inspection of the data? If not, perform appropriate analyses (perhaps both graphical and numerical) to try to identify the cause, then read the appropriate portion of Snee (1983) and compare your answer with that given by Snee. Do you agree with the conclusion given by the latter? Explain.

7.10 Gonzalez-de la Parra and Rodriguez-Loaiza (2003, references) gave an example of a nested design with three factors, to which they applied a proposed ANOM procedure that was described in Section 7.6. The objective of the study was to identify the primary source(s) of the overall variability of a synthetic process of a drug substance, as the operators believed that the quality of the product was related to the quality of the starting material. Two suppliers were randomly selected from the list of approved suppliers, three lots of starting material were selected at random from each supplier, and four containers from each lot were selected at random. Three assay determinations were then made on each container for the purpose of estimating the experimental error.

(a) There are some issues here relating to estimating the experimental error, including whether or not the same experimental unit is being measured.

Repeated measures designs are covered in Chapter 11, in which the same experimental unit is measured under different conditions. Is that the case here or, as discussed in Section 1.4.2, is it a matter of distinguishing between multiple readings and replications? In any event, based on the description of the study, do you believe that a reasonable estimate of the experimental error would result?

(b) What will be the degrees of freedom for containers?

7.11 Assume that an experiment in which a fully nested design was used had five factors, all of which were random. How would the divisor for the F-test be constructed for testing for the significance of the only factor that is not nested under another factor?

7.12 An example of a nested design that is often given is one that has been discussed throughout this chapter; namely, an experiment is performed to determine if there is any difference between machine performance in a manufacturing setting. Each machine has, say, four different heads (in four different positions on a machine) and it is felt that heads could influence the readings that are obtained from each machine. Each set of four heads can fit only a given machine, however, so heads are nested within machines. Hicks (1956, references) may have been the first to describe such an experiment with an illustrative example. Four observations were made on each head and the response variable was strain readings. The observations are given below.

	Machine																			
	1				2				3				4				5			
Head	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Readings	6	13	1	7	10	2	4	0	0	10	8	7	11	5	1	0	1	6	3	3
	2	3	10	4	9	1	1	3	0	11	5	2	0	10	8	8	4	7	0	7
	0	9	0	7	7	1	7	4	5	6	0	5	6	8	9	6	7	0	2	4
	8	8	6	9	12	10	9	1	5	7	7	4	4	3	4	5	9	3	2	0

Analyze the data and determine if there is a head effect or a machine effect (or both).

7.13 A survey is conducted in four classrooms in each of two schools, with each classroom–school combination thus constituting a group. Give the breakdown of the degrees of freedom for groups.

7.14 Liu and Batson (2003, references) described an interesting experiment that involved both control factors and noise factors, with the latter mentioned in Section 5.17.1. The experiment is somewhat unique in the sense that noise factors are generally regarded as being controllable under test conditions but not during normal plant operations, as is discussed in Chapter 8, but in this experiment they were not controllable. This means that the noise factors obviously could not be part of the experimental design. Therefore, regression analysis had to be used to determine if the noise factors were related to the response variable. That variable was “gauge gain,” which refers to the phenomenon of the gauge of a fabricated steel tube being thicker than that of the coil of steel that is used to produce the steel tube. The problem affects small-dimension tubes manufactured by cold-working.

The design was nested in the following way. There were three factors: gauge size (0.25 in., 0.188 in., and 0.12 in.), gauge shape (rectangular and square), and the size of tube for each shape in inches (3×4 and 2×5 for rectangular and 4×4 and 1.5×1.5 for square). From the latter, it is obvious that size of tube is nested under shape, this being the only nesting that occurs. Two replicates were used, so there were 24 observations. The data were as follows, with 11, 7, and 4 used below, in accordance with industry terminology, to represent the three gauge sizes given above; “1” denotes the rectangular shape and “-1” the square shape, with “1” denoting 3×4 and “2” denoting 2×5 for the rectangular shape, and “1” denoting 4×4 and “2” denoting 1.5×1.5 for the square shape.

Gauge Gain	Gauge	Shape	Size
0.004	11	1	1
0.004	11	1	1
0.002	7	1	1
0.003	7	1	1
0.004	4	1	1
0.003	4	1	1
0.004	11	1	2
0.005	11	1	2
0.010	7	1	2
0.012	7	1	2
0.007	4	1	2
0.011	4	1	2
-0.001	11	-1	1
0.000	11	-1	1
0.000	7	-1	1
0.001	7	-1	1
0.004	4	-1	1
0.002	4	-1	1
0.001	11	-1	2
0.002	11	-1	2
0.002	7	-1	2
0.000	7	-1	2
0.003	4	-1	2
0.001	4	-1	2

Analyze the data and determine significant effects, if any, recognizing that certain interactions are estimable even though there is nesting.

7.15 (Croarkin and Tobias, 2002, references) gave the following example in Section 3.2.3.3. Pin diameters are studied to see if there is any difference due to machines and operators. There are five machines that each have a day operator and a night operator. Five samples are taken from each machine and for each operator. The data, on pin diameters, are as follows.

	Machine				
	1	2	3	4	5
Operator (Day)	0.125	0.118	0.123	0.126	0.118
	0.127	0.122	0.125	0.128	0.129
	0.125	0.120	0.125	0.126	0.127
	0.126	0.124	0.124	0.127	0.120
	0.128	0.119	0.126	0.129	0.121
Operator (Night)	0.124	0.116	0.122	0.126	0.125
	0.128	0.125	0.121	0.129	0.123
	0.127	0.119	0.124	0.125	0.114
	0.120	0.125	0.126	0.130	0.124
	0.129	0.120	0.125	0.124	0.117

Analyze the data and draw appropriate conclusions.