Hasse Diagrams in Statistical Consulting and Teaching

Sharon L. LOHR

Hasse diagrams summarize the structure of mixed models and can be used by a statistical consultant to help design a complicated experiment or to help clarify the structure of data to be analyzed. They are also useful in the classroom as an aid for obtaining expected mean squares or deciding which denominator should be used in an F statistic.

KEY WORDS: Expected mean squares; Experiment design; Mixed model; Poset diagram.

1. INTRODUCTION

Statisticians often use a mixed model in their quest to obtain a simple structure to summarize a data set. The class of mixed models includes both fixed effects and random effects models; in the former the error term is the only random factor, and in the latter the overall mean is the only fixed factor.

Statistical package procedures such as SAS PROC GLM and BMDP-8V have considerably simplified the task of analyzing the data from a mixed-model ANOVA experiment. None of these programs, however, helps with specifying the model; a PROC GLM user who types in the wrong model will merely have faster incorrect results than if the calculations were done by hand.

Hasse diagrams, also known as poset or factor structure diagrams, can help in thinking about the structure of the experiment. They can also provide a quick check on the "reasonableness" of the computer results, sometimes giving clues that the model used in the analysis given by the computer is not the one that the data analyst thought was being used. They are useful in teaching others about design of experiments: because they are visual rather than analytic, they give students a supplementary perspective on different designs.

Hasse diagrams are widely used in combinatorics to display partially ordered sets. The combinatorial structure of factors in orthogonal ANOVA-type designs has long been known and exploited: see, for example, Taylor and Hilton (1981), Tjur (1984), Speed and Bailey (1987), and Andersson (1990). All of these authors used Hasse diagrams to display orthogonal designs; Taylor and Hilton

(1981) and Tjur (1984) illustrated how to use the diagrams to calculate sums of squares and degrees of freedom in orthogonal designs. Lindman (1992) displayed the hierarchical factor structure for some of his examples using Hasse diagrams.

In Section 2 of this paper the simple procedure for constructing the basic Hasse diagram for any ANOVA-type design is presented, assuming that the model structure is known. The basic Hasse diagram, showing the factor structure and degrees of freedom, can be constructed for nonorthogonal designs such as balanced incomplete block designs and factorials with unequal cell sizes, as well as for the orthogonal designs considered in the work cited above. This procedure can be taught in any data analysis class without having to discuss the underlying mathematics of partially ordered sets and distributive lattices; the reader interested in further exploration of the combinatorial structures is referred to Fishburn (1985).

Section 3 gives three examples of consulting sessions in which Hasse diagrams were used to help clarify the structure of the experiment. In Section 4 Hasse diagrams are used as a pedagogical aid to find sums of squares, expected mean squares, and the appropriate hypothesis test under normal theory for the null hypothesis that a given factor does not contribute to the variability of the data. Section 4 applies only to orthogonal mixed model designs such as full or fractional replicates of completely balanced factorial designs, randomized complete block designs, nested designs, Latin and Graeco–Latin squares, and balanced split-plot designs.

2. DRAWING THE DIAGRAM WHEN THE MODEL IS KNOWN

Consider a balanced two-way ANOVA experiment in which fixed factor A has two levels, random factor B has three levels, and there are two measurements at each level of A and B. A hypothetical set of data for such an experiment is shown in Figure 1a.

The model for this experiment contains five factors, corresponding to the five terms in the linear model

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \qquad (1)$$

where $i = 1, \ldots, a; j = 1, \ldots, b;$ and $k = 1, \ldots, n$. Besides the factors A and B and their interaction AB, there are the trivial factors M (for mean) and E (the error factor).

The partial order $<$ for a Hasse diagram comes from the nesting in a design. If factor F_2 is nested in factor F_1, then $F_2 < F_1$. In the Hasse diagram each factor is represented by a vertex. If $F_2 < F_1$, then F_2 is below F_1 and connected
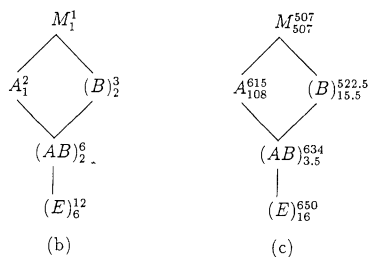
Figure 1. (a) Data, (b) Basic Hasse Diagram, and (c) Sums of Squares Calculations for the Sample Mixed Model Experiment Described in Section 2.

Figure 2. Hasse Diagram for Split-Plot Experiment in Snedecor and Cochran (1980, pp. 326–328).

to it by a path; if F_1 and F_2 are crossed factors, then no such one-way path exists. The partial ordering for the experiment described above is

$$E < AB < A < M, \quad E < AB < B < M.$$

Note that the interaction AB is nested in both A and in B, but factors A and B are crossed. The steps for constructing the basic Hasse diagram are given below, and illustrated in Figure 1b.

Step 1. Write down M at the top and E at the bottom. Thus the coarsest summary is at the top and the finest summary is at the bottom, just as in the ANOVA table. All other factors are then drawn between M and E. In a factorial design the main effects will be crossed on the first level, two-factor interactions on the second, three-factor interactions on the third, etc. In a completely nested design the factors will form a chain from top to bottom.

In the current example, A and B are crossed with each other but nested in M, so separate lines are drawn connecting M to A and M to B. Because AB is nested in both A and B, lines show that nesting as well. A final line connects E, which is nested within all the other factors, with AB.

Step 2. Put parentheses around the random factors. The error factor E is always random. Any factor between two random factors in a chain is automatically random. For the example in Figure 1 the interaction AB is random because it is between the random factors B and E.

Step 3. For each factor in the diagram, enter the number of factor levels as the superscript. In Figure 1b there is one level for the mean, six (the number of cells in the two-way table) levels for the AB interaction, and 12 (the number of observations) levels for E.

Step 4. Enter the degrees of freedom as the subscript for each level. Start at the top with one degree of freedom for the mean. Then work down through the other factors from top to bottom: for any factor F,

subscript of F = superscript of F

$\quad$ − (sum of all subscripts of factors above F). $\quad$ (2)
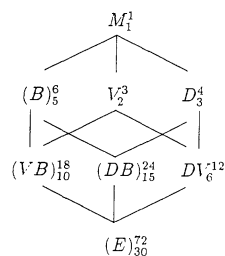
Steps 1–4 can be used to construct the basic Hasse diagram for any standard design, whether orthogonal or not. Figure 2 shows the diagram for a split-plot experiment described in Snedecor and Cochran (1980, pp. 326–328). The whole plot treatments (factor V) are three different varieties of alfalfa, and the subplot treatments (factor D) are four different dates of final cutting. The six replications (blocks) of the whole plots comprise factor B. The replicates and error term are both random, and so are placed in parentheses. In many split-plot experiments the factor corresponding to DB would be omitted from the diagram, so the three-factor interaction (here labeled as E) would be used as the subplot error. In this case, though, Snedecor and Cochran (1980, p. 327) note that, "since date of cutting effects are large in this experiment, we have reason to expect that the DB mean square will be larger than the VDB mean square, as was the case."

3. EXAMPLES: HASSE DIAGRAMS IN CONSULTING

The procedure in Section 2 for drawing Hasse diagrams assumes that the statistician knows the model before drawing the first line of the diagram. In a consulting session, though, the statistician often needs to extract the structure of the experiment as the investigator is describing it. Hasse diagrams allow the statistical consultant to keep track of various factors and their relationships as they are described, and to sketch out possible models on the spot.

The diagrams help the consultant with one of the biggest statistical problems investigators face: different sizes of experimental units. Many researchers have learned about cluster samples or split-plot designs in classes, but do not recognize their own experiment as such. Hurlbert (1984, p. 208) stated that "Pseudoreplication is probably the single most common fault in the design and analysis of ecological field experiments," and his statement applies to many other disciplines as well. The consultant is often able to see a split-plot or repeated-measure structure or nesting more easily by using a Hasse diagram than by taking notes with words, by trying to write down the mathematical model during the session, or by simply trying to remember everything.

The investigator will often ask about the Hasse diagrams while the statistician is taking notes; this provides an opportunity to explain some of the replication issues in an easy-to-understand graphical manner without using subscripted Greek letters or having to provide a label for the design. In this respect the Hasse diagrams fit in well

with the no-name approach advocated by Lorenzen and Anderson (1993).

The remainder of this section gives three examples of consulting sessions in which the author used Hasse diagrams. The conversations have been edited for brevity.

Example 1. Investigator: I'm writing a dissertation on teacher workload in Arizona, to see what the actual workload is and whether it can be predicted by some of the teacher workload formulas that have been developed.

Statistician: What do you mean by workload: the number of classes taught, the number of hours worked, or what?

Investigator: Workload is the number of hours a teacher spends on school-related activities in a typical week. The formulas that have been developed predict workload from number of classes taught, extracurricular activities, and other factors. I want to take a survey of elementary school teachers to see how much they actually work and to validate the formulas.

Statistician: How were you thinking about selecting teachers for your sample?

Investigator: Well, I had thought about just taking a random sample of teachers in the state, but there is no list of all the teachers. But once I contact a school, the site administrator can give me a list of all the teachers.

Statistician: So you want to take a sample of schools? Do you want to be able to compare workloads for different areas of the state? [Draws Fig. 3a.]

Investigator: Actually, I want to compare workloads for schools in large, medium, and small school districts. The district size is defined by the average daily student membership. The largest districts are in urban areas and the smallest in rural, so I will be able to look at the difference between urban and rural areas.

Statistician: Does that mean you would need to look at schools by district?

Investigator: Yes, it would. I suspect that workload will be different in the different sizes of school districts.

Statistician: Let's talk a bit about how you're going to choose the schools to be in your sample.

Investigator: Well, most studies select schools using a systematic sample, and I'd like to follow precedent. I want to stratify schools by the total daily membership of the district. Then I'd like to list the schools within each school district for each stratum. Then I'd like to take a 10% sample.
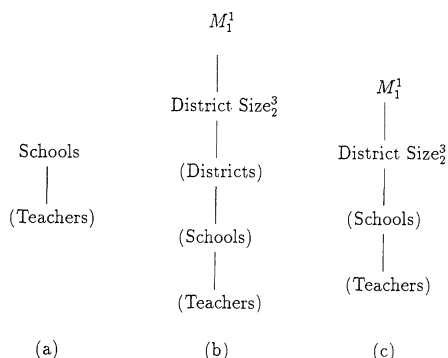
Statistician: So you have strata, and then you're selecting districts from each stratum, schools from each district, and then all the teachers from each school? [Draws Fig. 3b.]

Investigator: I don't think that's quite it. I'm using district size for the strata, and then taking a subsample of schools from each of the three groups of schools.

Statistician: [Draws Fig. 3c.] Let's talk a little about what your sample sizes need to be, and how you're going to get a good response rate . . . .

Example 2. Investigator: I'm studying the effects of feral pigs on the native vegetation in Santa Cruz Island, California [gives background of the island; see Peart (1994)].

Statistician: What do your data look like?

Investigator: I've been following areas around ten oak trees on the island since 1989. Under the canopy of each tree, volunteers from The Nature Conservancy built 36-square-meter exclosures that keep pigs out but let all of the other mammals and birds of the island in. Next to each exclosure, we staked out an area of equal size, but did not put up any barriers. This is what makes SCI ideal for this study—if we did it on the mainland, we'd have to worry about whether the damage was caused by feral pigs, deer, sheep, cattle, or pocket gophers. Of those animals, only pigs wander around on SCI. Every three to six months I visit all ten sites and record the amount and location of seedlings and other vegetation, and also measure the amount of pig activity.

Statistician: So you have two plots for each tree: one fenced and one unfenced? [Draws Fig. 4a.]

Investigator: No, I have four plots: I have another pair of plots, one fenced and one unfenced, about 15 feet away from each tree not under its canopy.

Statistician: Let me see if I'm understanding you correctly. Is this a typical tree with its plots? [Draws Fig. 4b.]
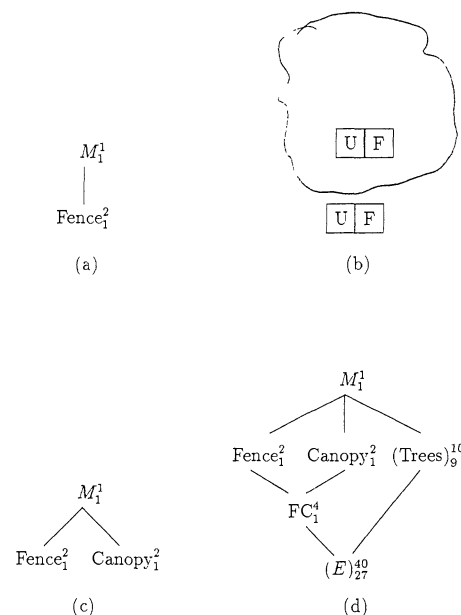

(a) (b)

Figure 4. Hasse Diagrams Used in Consulting Session for Oak Seedling Study.

(a)   (b)   (c)

Figure 3. Hasse Diagrams Used in Consulting Session for Teacher Workload Study.

(c)   (d)

Investigator: Yes, except I don't always have them lined up like that, and the control plots can be on either side of the exclosures.

Statistician: [Draws Fig. 4c.] When you chose the individual trees for the study, did you choose them because of specific properties such as this one's close to water, this one's at a higher elevation, or did you just want trees that are representative of all oaks on the island?

Investigator: I think I have trees that are representative of the oaks on the island.

Statistician: [Draws Fig. 4d.] What are you trying to find out from the data?

Investigator: I want to see whether the pigs are inhibiting woody seedling germination, development, and survival, and whether they are affecting the mix of native and exotic plants. I've run some preliminary t tests on the data under the canopy, and the results are overwhelmingly significant.

Statistician: These p values certainly do look significant, but I think that may be because the computer package you're using doesn't distinguish between different types of experimental units. One of the basic assumptions of the t test is that all the observations are independent, and that doesn't happen for your data since successive measurements on the same tree are likely to be more similar than measurements on different trees. If we look schematically at your design, you really have ten experimental units, the ten trees. Within each tree you're looking at the effect of the fences and the effect of being under the canopy, and you also may have an interaction—keeping pigs out may have more of an effect under the canopy than away from the tree. So if you were running an ANOVA on some summary measure such as average number of seedlings over time, you would only have 27 degrees of freedom for the error term.

Example 3. Investigator: This study is designed to answer the question: "Does the period of wearing time have an effect on the tests we use to fit hearing aids?" As people get used to wearing a hearing aid, data show that there may be an improvement in their scores on speech understanding because people learn to recognize speech cues that they had previously not been hearing. Not all researchers see such a long-term effect, however. I want to design a new study, using each person as his or her own control. Knowing my population, I could come up with 30 people who could complete a series of sessions.

Statistician: What kind of study do you have in mind?

Investigator: I'd like to get people who have never worn a hearing aid before and test their speech understanding. They come back two days later, and we repeat the measurements so I can get an idea of the variability of each individual. Then I send them off with a hearing aid in each ear and have them repeat the measurements over a period of time to look for a possible acclimatization effect.

Statistician: Why have everyone start without a hearing aid and then wear them in both ears? Why not send each person home with one hearing aid to get used to, and then be able to compare the ear that has been wearing the

hearing aid with the ear that has not been wearing a hearing aid? It seems that then you would eliminate much of the variability between subjects.

Investigator: The problem is that it is very difficult to find people who have similar hearing losses in each ear. A previous researcher did what you're describing, and only found four subjects who qualified for the study. I also want to study two types of hearing aid, to see if one has a greater effect than the other. I thought about having each subject wear both types of hearing aids, first one type, then the other in random order, but decided that would make the experiment too long. So I thought I'd use 15 people for each type of hearing aid.

Statistician: So you have two different types of hearing aids, and each subject is only going to be wearing one type of hearing aid, but wearing that type of aid in each ear?

Investigator: Right.

Statistician: Are you interested in the individual subjects per se? Are you deliberately picking them so that subject A has a specific hearing loss and subject B has a different hearing loss, etc.?

Investigator: I'm going to have a reasonably homogeneous group, and randomly assign half to each hearing aid.

Statistician: [Draws Fig. 5a.] Here I've drawn type of hearing aid, and I've put subjects below the type because any subject belongs completely to the group using that one type of hearing aid.

Investigator: And this is the number of subjects?

Statistician: That's the number of degrees of freedom for subjects.

Investigator: There's one more twist in that I'm letting people adjust the volume level. I want to do two tests: one in which the volume is fixed, and the other in which I allow the patient to adjust the volume. So in effect we have three treatment groups: the original measurement, with fixed volume, and with adjustable volume.

Statistician: [Draws Fig. 5b.] But aren't you looking at the difference between the measurement with the hearing aid and that without as your response? [Draws Fig. 5c as a possible alternative to 5b.] I'm a little worried here about the possibility that, with the different volume settings done on the same patient, you may get a carryover effect. How will you know in this experiment whether you're getting an acclimatization to the hearing aid or whether people are getting better at the speech test?

Investigator: Well, I was concerned about that, too. I'd like to randomize the order in which the volume settings
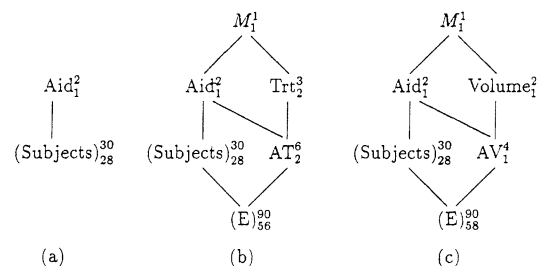

Figure 5. Hasse Diagram Used in Consulting Session for Hearing Aid Study.

are done to try to minimize that problem. But I don't see how to look for acclimatization without giving people the same basic speech test over a period of time. And that will give me an idea about the variability in our speech testing.

Statistician: I think this sounds like a really interesting study. You might want to get some data on the reliability of your speech tests before doing the rest of the study. You could do that using your current patients or patients who don't qualify for your new study.

4. SUMS OF SQUARES, EXPECTED MEAN SQUARES, AND HYPOTHESIS TESTS FOR ORTHOGONAL DESIGNS

The most useful parts of Hasse diagrams for consulting sessions were described in Section 2 and demonstrated in Section 3. Hasse diagrams can also be used as a tool for teaching other calculations in a design course. Most calculations that can be done using Hasse diagrams are also found in various statistical software packages, but the Hasse diagrams provide an additional visual perspective that many students find helpful. Note that, although the basic Hasse diagram can be drawn for any design, simple algorithms for calculating sums of squares and expected mean squares only exist for orthogonal designs. Unbalanced data require that some effects be adjusted for the contribution of other effects and such adjustments are best done using the computer.

For orthogonal designs, sums of squares may be computed in much the same way as degrees of freedom. For factor F let

$$S_F = \sum_{\text{levels } f \text{ of factor } F}$$
$$\times \frac{(\text{sum of all observations at level } f \text{ of factor } F)^2}{(\text{number of observations at level } f \text{ of factor } F)}.$$

Thus

$$S_M = \frac{\text{sum of all observations})^2}{(\text{number of observations})}$$

and

$$S_E = \text{sum of squares of all observations.}$$

In a sums-of-squares Hasse diagram, S_F is written as the superscript of factor F, S_M is written as both the superscript and the subscript of factor M, and the subscripts of the other factors are found using (2). The subscripts are then the sums of squares. Such sums of squares calculations are illustrated in Figure 1c for the hypothetical example of Section 2.

Finding expected mean squares and the appropriate test statistic is a little more complicated, as the expected mean squares depend on the particular parameterization of the mixed model used. Two parameterizations in common use are described in Schwarz (1993) as Formulation 1 and Formulation 2. For the mixed model in (1), Formulation 1 assumes that

1(a) α_i are parameters
1(b) β_j are iid $N(0, \sigma_B^2)$
1(c) ϵ_{ijk} are iid $N(0, \sigma^2)$

1(d) $(\alpha\beta)_{ij}$ are iid $N(0, \sigma_{AB}^2)$
1(e) β_j, $(\alpha\beta)_{ij}$, and ϵ_{ijk} are independent for all i, j, and k.

The assumptions for Formulation 2 are similar, except that interaction terms involving at least one fixed factor sum to zero over the fixed factor indices. For the model in (1) the assumptions for Formulation 2 are

2(a) α_i are parameters with $\sum \alpha_i = 0$
2(b) β_j are iid $N(0, \xi_B^2)$
2(c) ϵ_{ijk} are iid $N(0, \xi^2)$
2(d) $(\alpha\beta)_{ij}$ are $N(0, \frac{a-1}{a}\xi_{AB}^2)$ with $\sum_i(\alpha\beta)_{ij} = 0$ for all j and cov $((\alpha\beta)_{ij}, (\alpha\beta)_{rj}) = (-\xi_{AB}^2/a)$ for $i \neq r$
2(e) β_j, $(\alpha\beta)_{ij}$, and ϵ_{ijk} are pairwise independent.

The choice of formulation depends on the covariance structure of the data. Here, procedures for finding expected mean squares are presented for both formulations. The Hasse diagram allows the expected mean squares to be calculated without constructing additional tables and without keeping track of "live" and "dead" subscripts as in Taylor and Hilton (1981).

For either formulation it is easiest to start at the bottom of the basic Hasse diagram and work upward. If F is a fixed factor, let $Q(F)$ be a quadratic function involving the treatment-effect parameters for fixed factors containing F. For factor A in model (1), for example, $Q(A) = \sum_{i=1}^{a}(\alpha_i - \alpha_.)^2/(a - 1)$. The following two steps give the expected mean square for any factor F in an orthogonal design, using Formulation 1.

Step 1. If F is random, then write down

$$\frac{\text{total number of observations}}{\text{superscript on } F} \times \sigma_F^2.$$

If F is fixed, then write down

$$\frac{\text{total number of observations}}{\text{superscript on } F} \times Q(F).$$

Step 2. For every random factor R below F, include the term

$$\frac{\text{total number of observations}}{\text{superscript on } R} \times \sigma_R^2.$$

The steps for Formulation 2 are similar, except that ξ_F^2 replaces σ_F^2 in Step 1, and Step 2 is replaced by Step 2':

Step 2'. For every random factor R below F, include the term

$$\frac{\text{total number of observations}}{\text{superscript on } R} \times \xi_R^2$$

if every fixed factor found in R is also in F.

This procedure is used in Figure 6a to find expected mean squares for the experiment of Figure 1 under the assumptions of Formulation 1. Since the error term is a random effect with nothing below it, the expected mean square is $(12/12) \times \sigma^2$. The interaction AB is also random. Step 1 says to write down $(12/6)\sigma^2$; Step 2 says to add σ^2 to that. For random factor B the expected mean square has term $(12/3)\sigma_B^2$ from Step 1, and $2\sigma_{AB}^2 + \sigma^2$ from Step 2.

The expected mean squares calculations under Formulation 2, shown in Figure 6b, are the same except for
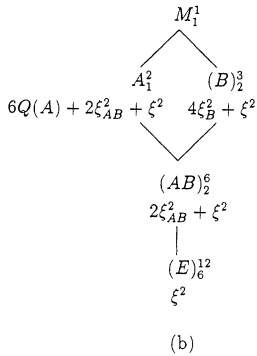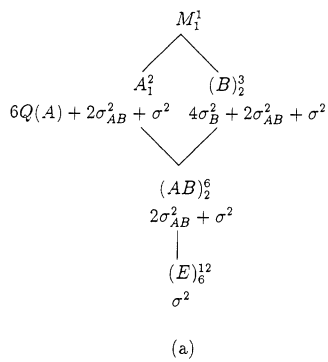
(a)

(b)

Figure 6. Expected Mean Squares for the Experiment in Figure 1 Calculated Using the Assumptions in (a) Formulation 1 and (b) Formulation 2.

(a) Formulation 1

(b) Formulation 2

Figure 7. Expected Mean Squares for Three-Factor Model.

$E[MSB]$. Since the interaction AB contains the fixed factor A not found in factor B, the term ξ_{AB}^2 does not appear in $E[MSB]$.

Figure 7 shows the expected mean squares under the two sets of assumptions for a completely balanced experiment with fixed factors A and B and random factor C. Factors A, B, and C have a, b, and c levels, respectively, and there are n replications at each factor combination.

The expected mean squares calculations lead directly to simple rules for which mean square to use in the denominator in the F statistic for testing the null hypothesis that the factor does not contribute to the variability of the data.

Formulation 1. The denominator is the mean square of the closest random factor below the effect to be tested, provided there is just one.

Formulation 2. The denominator is the mean square of the closest random factor below the factor to be tested that does not include a new fixed effect. If more than one factor qualifies as the closest, an approximation is needed.

The expected mean squares need not be calculated explicitly to be able to set up the F tests. The diagram in Figure 1b can be used directly to set up the appropriate F statistics. The F statistics for testing the interaction AB and for the main effect of factor A are the same for both formulations: MSAB/MSE and MSA/MSAB, respectively. For testing the effect of B, Formulation 1 uses MSB/MSAB because AB is the nearest random factor below B; Formulation 2 uses MSB/MSE for the F statistic because the closer random effect AB contains fixed effect A.
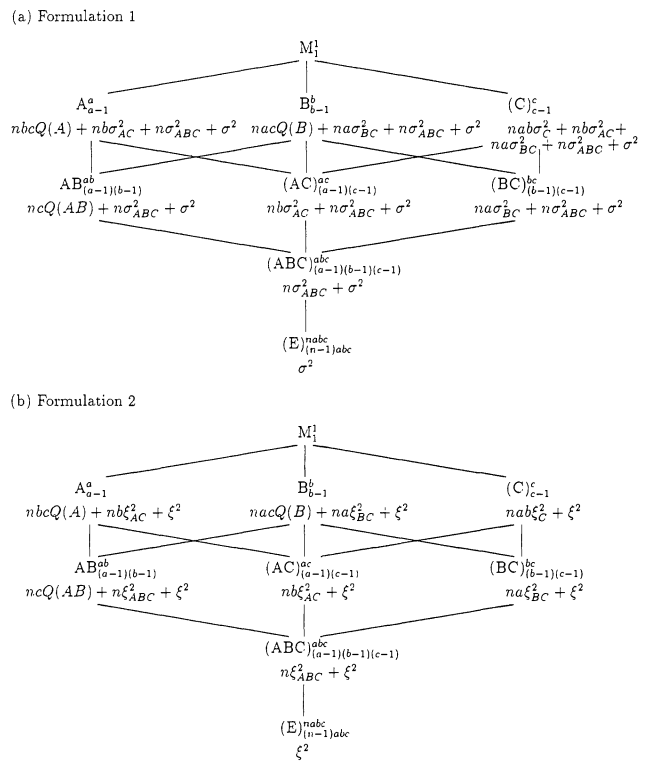
The diagrams also indicate when an exact F test cannot be found in a mixed multifactor model. In Figure 7a, for example, AC and BC are two different "closest" random factors below factor C, and no exact F test exists. The expected mean squares may, however, be used to construct a Satterthwaite approximate F test.

REFERENCES

Andersson, S. A. (1990), "The Lattice Structure of Orthogonal Linear Models and Orthogonal Variance Component Models," *Scandinavian Journal of Statistics*, 17, 287–319.

Fishburn, P. C. (1985), *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets,* New York: John Wiley.

Hurlbert, S. H. (1984), "Pseudoreplication and the Design of Ecological Field Experiments," *Ecological Monographs*, 54, 187–211.

Lindman, H. R. (1992), *Analysis of Variance in Experimental Design,* New York: Springer-Verlag.

Lorenzen, T. J., and Anderson, V. L. (1993), *Design of Experiments: A No-Name Approach,* New York: Marcel Dekker.

Peart, D. (1994), "Impacts of Feral Pig Activity on Vegetation Patterns Associated with *Quercus Agrifolia* on Santa Cruz Island, California," unpublished Ph.D. dissertation, Arizona State University.

Schwarz, C. J. (1993), "The Mixed-Model ANOVA: The Truth, the Computer Packages, the Books," *The American Statistician*, 47, 48–59.

Snedecor, G. W., and Cochran, W. G. (1980), *Statistical Methods* (7th ed.), Ames, IA: The Iowa State University Press.

Speed, T. P., and Bailey, R. A. (1987), "Factorial Dispersion Models," *International Statistical Review*, 55, 261–277.

Taylor, W. H., and Hilton, H. G. (1981), "A Structure Diagram Symbolization for Analysis of Variance," *The American Statistician*, 35, 85–93.

Tjur, T. (1984), "Analysis of Variance Models in Orthogonal Designs" (with discussion), *International Statistical Review*, 52, 33–81.