Simplified diagrams, models, and expected mean squares

Russell V. Lenth

STAT:5201 Experimental Design

Important note The material in this handout *replaces* the material in sections 12.6 and 12.7. I expect you to learn the methods in this handout, and to use these methods *instead* of Hasse diagrams in homework and exams.

1 Introduction

Sections 12.6 and 12.7 present methods based on Hasse diagrams. Those methods are correct and they can be helpful; but as the number of factors grows, they become cumbersome. In addition, Hasse diagrams are based on the idea that you include *every valid term* in the model. In many experiments (especially those with many factors) you may want to consider omitting some terms.

This handout provides methods based on the simple diagrams discussed in Section 2.4 of *DDD*, which show only the *factors* in the experiment, rather than all the model *terms*. I like these better because the diagrams are simple and they give you a clearer and less cluttered picture of the structure of the experiment.

Notational change Besides using different diagrams, I also depart from the text in this way: fixed factors are denoted in lower case and random factors are UPPER CASE.

Example 1 (Laundry experiment) This experiment is designed to investigate shrinkage of fabrics when dried in an automatic dryer. Four different fabrics (cotton, rayon, polyester, and wool) are to be compared at three different dryer temperatures (low, medium, high). Fixed-size squares of material are first washed, then dried at a selected temperature until it automatically shuts off. The dimensions of the squares are measured, and that forms the response variable. We will run 5 dryer loads at each temperature; each load will have one square of each fabric. Thus, there are $N = 4 \times 3 \times 5 = 60$ measurements in the experiment.

This experiment has three factors: temp, fabric, and LOAD. The latter is taken to be random because we'd expect there to be random variations from one load to the next. Each load has only one temperature, but all 4 fabrics. That is, LOAD is nested in temp and crossed with fabric. The simplified diagram for this experiment is this:

$$\text{temp}^3 \quad \times \quad \text{fabric}^4$$

$$|$$

$${}^5\text{LOAD}^{15}$$

We use superscripts to show the number of levels (as in the textbook's Hasse diagrams). In nested factors, a left superscript (not in the textbook) shows the number of levels *per* level of the factor above it (in this case, 5 loads per temperature), and the right superscript shows the total number of levels (15 loads altogether).

2 Developing a model

The model will usually contain all of the main factor effects, which have the same symbols as the diagram. We will write models SAS-style, using * to denote interactions and parentheses to denote nesting: A(b) denotes A nested in b. For the laundry example, the main effects are temp, LOAD(temp), and fabric.

In addition, we can include interactions among factors *provided they don't lie on the same nesting chain in the diagram.* In the laundry example, the legal interactions are temp*fabric and fabric*LOAD(temp) (the latter corresponds to individual observations in the experiment, so usually we will leave it out and call it RESIDUAL instead). An interaction between temp and LOAD(temp) is *not* legal, as they lie in the same vertical chain (a hint that this interaction is illegal is that they both involve the same factor, temp). When an interaction involves a nested factor, SAS requires us to put the parenthesized part at the end; if it involves two nested factors, we combine the nesting factors in one set of parentheses.

3 Degrees of freedom

To obtain the d.f. for a term, consider each factor involved in the term, and write down its number of levels (superscript, or *left* superscript if a nested factor). Subtract 1 if that factor is *not* inside parentheses. Then take the product of all these results.

Example 2 (Laundry experiment) Here are the degrees of freedom for our example:

Term	levels	df
temp	3	$3-1=2$
LOAD(temp)	5	$(5-1)\times 3=12$
fabric	4	$4-1=3$
temp $*$ fabric	3, 4	$(3-1)\times(4-1)=6$
fabric $*$ LOAD(temp)	4, 5, 3	$(4-1)\times(5-1)\times 3=36$

Note that these sum to 59, which is equal to $N-1$; thus, all degrees of freedom are accounted for.

You may follow the textbook's convention of showing degrees of freedom as subscripts in the diagram:

$$\begin{array}{ccc} \text{temp}^{3}_{2} & \times & \text{fabric}^{4}_{3} \\ | & & \\ {}^{5}_{4}\text{LOAD}^{15}_{12} & & \end{array}$$

The d.f. for an interaction is the product of the d.f. for the main effects it comprises. In the diagram, note that both right-hand scripts on LOAD can be obtained my multiplying the left scripts by the right superscript (in this case, 3) of the factor above. This will always be true.

4 Expected mean squares

Unrestricted case only We give rules only for the "unrestricted model"—and that is the only case we will use in this course. In reality, both the restricted and unrestricted "models" are different parameterizations of the same model; so there is really little value in doing both.

As in the textbook, every model term t that consists *entirely* of fixed effects has a corresponding fixed variation $Q(t)$ that is equal to the sum of the squares of all its effect values (including repeats), divided by the degrees of freedom for t. Every other term involves at least one random factor, and there is a corresponding variance component σ^2_t for the variance of those effects.

If the degrees of freedom for the terms in the model sum to less than $N-1$, be sure to include a term for RESIDUAL that accounts for these extra degrees of freedom.

You can obtain the expected mean squares in two steps:

1. Find each term t's "leading component": If a fixed effect, this is just $Q(t)$; if it is a random effect, it is σ^2_t multiplied by the number of times each level of that effect is repeated. (For terms that appear in the diagram, the multiplier is N divided by the right-hand superscript; for interactions, divide N by each of the superscripts involved.)

2. For each term t, add the leading components of all other *random* or *mixed* terms that contain every factor in t. (Note: RESIDUAL contains everything.)

Note that a $Q()$ quantity can appear only as the leading component of a fixed effect's EMS.

Example 3 (Laundry) Here are the results of step 1, showing only the leading components:

Source	df	EMS	
temp	2	$Q(t)$	$+\cdots$
LOAD(temp)	12	$4\sigma^2_{L(t)}$	$+\cdots$
fabric	3	$Q(f)$	$+\cdots$
temp $*$ fabric	6	$Q(tf)$	$+\cdots$
fabric $*$ LOAD(temp)	36	$\sigma^2_{fL(t)}$	$+\cdots$

The coefficient of $\sigma^2_{L(t)}$ is $4=60/15$ reflects the fact that there are 4 observations in each load. Mechanically, the coefficient of $\sigma^2_{fL(t)}$ is $60/(4\times 15)=1$, reinforcing the point that this term reflects individual observations: there is only one observation at each combination of fabric and load within temperature.

To complete step 2, we simply add-in the leading components of random effects that entirely contain the one in question. For example, the terms containing temp are LOAD(temp), temp*fabric, and fabric*LOAD(temp); however, temp*fabric is a fixed effect, so only the first and third of these contributes to the EMS for temp. The complete EMS table is

Source	df	EMS		
temp	2	$Q(t)$	$+4\sigma^2_{L(t)}$	$+\sigma^2$
LOAD(temp)	12		$4\sigma^2_{L(t)}$	$+\sigma^2$
fabric	3	$Q(f)$		$+\sigma^2$
temp $*$ fabric	6	$Q(tf)$		$+\sigma^2$
RESID $=$ fabric*LOAD(temp)	36			σ^2

I got lazy and wrote σ^2 in place of $\sigma^2_{fL(t)}$. Note that to improve readability, it is advisable to add a new alignment point for each new random effect, as shown. I also added a horizontal rule to emphasize the distinction between the "between-load" and "within-load" effects.

5 Means, comparisons, contrasts

To estimate a mean or some comparison or contrast thereof, the main task is to figure out how to estimate the variance of the quantity of interest. Then we can construct CIs or tests using the usual normal-theory techniques. For example, if $\hat\theta$ is an estimator of some linear function θ of true model effects, and $\mathrm{Var}(\hat\theta)=\omega$, then a $(1-\mathcal{E})$ CI for θ is $\hat\theta \pm t_{\mathcal{E}/2,\nu}\sqrt{\hat\omega}$ where $\hat\omega$ is an estimate of ω having ν d.f.

To understand how to do this, consider a formal model for the Laundry example:

$$y_{ijk}=\mu+\tau_i+L_{j(i)}+\phi_k+\tau\phi_{ik}+\epsilon_{kj(i)}$$

where i,j,k are the subscripts for temp, LOAD, and fabric, respectively. The ith temp mean is thus

$$\bar{y}_{i\bullet\bullet}=\mu+\tau_i+\bar{L}_{\bullet(i)}+\bar{\phi}_{\bullet}+\overline{\tau\phi}_{i\bullet}+\bar{\epsilon}_{\bullet\bullet(i)}$$

Only the *random* terms here contribute to the variance; since (in the unrestricted model) the terms are independent, it follows that

$$\text{Var}(\bar{y}_{i\bullet\bullet}) = \text{Var}(\bar{L}_{\bullet(i)}) + \text{Var}(\bar{\epsilon}_{\bullet\bullet(i)})$$

Now, for each i, $\bar{L}_{\bullet(i)}$ is the average of the 5 independent LOAD effects for that temperature; so $\text{Var}(\bar{L}_{\bullet(i)}) = \sigma^2_{L(t)}/5$. Similarly, there are 20 independent f*L(t) effects at each temperature, so $\text{Var}(\bar{\epsilon}_{\bullet\bullet(i)}) = \sigma^2/20$. It follows that $\text{Var}(\bar{y}_{i\bullet\bullet}) = \sigma^2_{L(t)}/5 + \sigma^2/20$. An unbiased estimate of this quantity is $MS_{L(t)}/20$ (refer to the EMS table), with 12 d.f.; and a 95% CI for $\mu_{i\bullet\bullet}$ is

$$\bar{y}_{i\bullet\bullet} \pm t_{.025,12}\sqrt{MS_{L(t)}/20}$$

Similarly, if we are computing a contrast of temperature means

$$se(\mathbf{w}(\{\bar{y}_{i\bullet\bullet}\})) = \sqrt{(MS_{L(t)}/20)\cdot\sum w_i^2}$$

In the special case of comparing two means, $\mathbf{w}$ includes a 1, a -1, and the rest are zero, so that $\sum w_i^2 = 2$ and this standard error is $\sqrt{MS_{L(t)}/10}$.

The above approach can be used successfully to figure out variances for other means and contrasts, and in other experiments. But here's another notation that you may find friendlier. First of all, we need only to keep track of the random and mixed effects—LOAD(temp) and fabric*LOAD(temp) in the laundry example. To save writing, let's call them just L(t) and f*L(t) from now on. For a temp mean, the contributions of these effects can be denoted

$$\mathsf{L}^\bullet(\mathsf{t}) \quad \text{and} \quad \mathsf{f}^\bullet{*}\mathsf{L}^\bullet(\mathsf{t})$$

to represent the dots in the subscripts of the formal model. As before, we then observe that there are 5 loads and 20 f*L(t) effects in each temp mean, so these are the divisors for their respective variances.

In general, it is helpful to name every factor in the residual error component (including REP for replications, if applicable). The dots tell us which superscripts in the diagram (use the *left* ones in nested factors) to multiply together to obtain the divisor. For the temp means, $\mathsf{L}^\bullet(\mathsf{t})$ gets a divisor of 5, the superscript for LOAD, and $\mathsf{f}^\bullet{*}\mathsf{L}^\bullet(\mathsf{t})$ gets a divisor of $4 \times 5 = 20$.

Let's do the other means using this approach. For the fabric means, the contributions of the random effects are

$$\mathsf{L}^\bullet(\mathsf{t}^\bullet) \quad \text{and} \quad \mathsf{f}{*}\mathsf{L}^\bullet(\mathsf{t}^\bullet)$$

Each mean involves all 15 L(t) effects, and 15 f*L(t) effects; so the variance of a fabric mean is

$$\text{Var}(\text{fabric mean}) = (\sigma^2_{L(t)} + \sigma^2)/15$$

This must be estimated using a combination of mean squares, and a Satterthwaite d.f. formula. An interesting thing happens when we compare or contrast the fabric means though: Given contrast coefficients $\mathbf{w}$, $\mathbf{w}(\{\mathsf{L}^\bullet(\mathsf{t}^\bullet)\})$ reduces to zero, because it is constant across fabrics; thus, the variance of the contrast is simply

$$\text{Var}(\mathbf{w}(\{\text{fabric}\})) = (\sigma^2/15)\cdot\sum w_f^2$$

which can be estimated using $(MS_E/15)\cdot\sum_f w_f^2$. These results make sense intuitively: the variations among dryer loads contribute to the uncertainty of estimating fabric means. However, when comparing or contrasting these means, these are within-LOAD contrasts and thus the variations among loads don't play a role.

Finally, consider the mean at each combination of temp and fabric. The random quantities involved are

$$\mathsf{L}^\bullet(\mathsf{t}) \quad \text{and} \quad \mathsf{f}{*}\mathsf{L}^\bullet(\mathsf{t})$$

These means involve 5 loads and 5 f*L(t) effects; thus,

$$\text{Var}(\text{temp*fabric mean}) = (\sigma^2_{L(t)} + \sigma^2)/5$$

If we want to estimate a contrast of these cell means, think of $\mathbf{w} = \{w_{tf}\}$ as being doubly subscripted because we need to keep track of temperature and fabric combinations. The $\mathsf{L}^\bullet(\mathsf{t})$ effects are constant for all fabrics, so what matters there are the marginal contrast coefficients $w_{t+} = \sum_f w_{tf}$. The variance of the contrast is

$$\text{Var}(\mathbf{w}(\{\text{t*f}\})) = \frac{\sigma^2_{L(t)}}{5}\cdot\sum_t w_{t+}^2 + \frac{\sigma^2}{5}\cdot\sum_t\sum_f w_{tf}^2$$

Most commonly, we want to do pairwise comparisons within a row or column of the table of cell means, so that one $w_{tf} = 1$, one $w_{tf} = -1$, and the rest are zero. Note that if the comparison is within a row (temp is held fixed), the w_{t+} values are all zero, whereas if it is within a column (fabric is held fixed), we obtain a $w_{t+} = 1$ and another $w_{t+} = -1$. Thus,

$$\text{Var}(\text{t*f diff}) = \begin{cases} 2\sigma^2/5 & \text{same temp} \\ 2(\sigma^2_{L(t)} + \sigma^2)/5 & \text{otherwise} \end{cases}$$

Referring to the EMS table, we can estimate the first quantity using $2 \times MS_E/5$, with 36 d.f. For comparisons involving different temperatures,

$$\widehat{\text{Var}} = \frac{MS_{L(t)} + 3 \times MS_E}{10}$$

$$df = \frac{(MS_{L(t)} + 3 \times MS_E)^2}{\frac{MS^2_{L(t)}}{12} + \frac{9 \times MS_E^2}{36}}$$

Again, these results make intuitive sense: comparisons at the same temperature are within-LOAD comparisons, whereas comparisons at different temperatures are between-LOAD, and hence more variable.