# 1 Analysis for Factorial Treatment Designs

Understanding the results from analyses of factorial treatment designs is aided by recalling the types of contrasts being tested. Sometimes it is easiest to think in terms of the cell means model:

$$y_{ijk} = \mu_{ij} + e_{ijk},$$

while other times it is more convenient to use the effects model:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk}.$$

The models in this chapter all involve fixed effects, so the effects model has the restrictions that:

$$\sum_{i=1}^{a} \alpha_i = 0, \sum_{j=1}^{b} \beta_j = 0, \text{ and } \sum_{i=1}^{a} \alpha\beta_{ij} = \sum_{j=1}^{b} \alpha\beta_{ij} = 0.$$

Since Proc GLM in SAS understands the notation of the effects model, it is often easiest to think of contrasts in terms of the means model, then translate into the effects notation using the relationship:

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij}.$$

We will illustrate these concepts by looking at the example from the text, first using the SLICE option in the LSMEANS statement, then by using the CONTRAST and ESTIMATE statements to calculate sums of squares and estimates directly.

## 1.1 Standard errors and interval estimates for means

## 1.2 Extending to the case of three factors

## 1.3 Power and Sample Size for Factorial Treatment Designs

The expressions for the noncentrality parameters for each of the three tests (main effect A, main effect B, interaction of A and B) are presented in section 6.8, and a SAS macro for computing sample sizes for main effects is on the website.

## 1.4 Addressing Unbalanced Data for Factorial Treatment Designs

Thus far in the chapter we have assumed that we had balanced factorial data, meaning that there were the same number of replicates for each combination of treatment factors. When we have unbalanced factorial data, some issues arise as to how to conduct tests of hypotheses. Recall in section 2.10 of the text, where the concept of defining a full and a reduced model was presented as a general way to conduct tests with linear models. With unbalanced factorial data, our sum of squares for a factor can depend upon what models are defined as the full and reduced models for the test. We will use the notation that $SS(A|B)$ is the model sum of squares to test the effect of $A$ while 'adjusting' for effect $B$. Thus $SS(A|B) = SS(AB) - SS(B)$, where $SS(AB)$ is the model sum of squares for the model with both the $A$ and $B$ effects. For a two-factor model with interaction, we will use the letter $C$ to denote the interaction effect, to avoid confusion with the $AB$ model above. With this notation we can explain the difference between SAS Type I, II, and III sums of squares. Suppose the model is written to SAS as: y = A B AB, then the Type I SS are formed as:

$$\begin{aligned} SS(A) &= SS(A|\oslash) \\ SS(B|A) &= SS(AB) - SS(A) \\ SS(C|AB) &= SS(ABC) - SS(AB) \end{aligned}$$

The Type II SS are formed as:

$$\begin{aligned} SS(A|B) &= SS(AB) - SS(B) \\ SS(B|A) &= SS(AB) - SS(A) \\ SS(C|AB) &= SS(ABC) - SS(AB) \end{aligned}$$

The Type III SS are formed as:

$$\begin{aligned} SS(A|BC) &= SS(ABC) - SS(BC) \\ SS(B|AC) &= SS(ABC) - SS(AC) \\ SS(C|AB) &= SS(ABC) - SS(AB) \end{aligned}$$

The pattern that emerges is that Type I SS are sequential, Type III SS adjust for every term in the model, and Type II SS adjust for all terms at the same level or below the term being tested. There is an active debate about whether Type II SS or Type III SS are more appropriate for tests with unbalanced data, but most would agree that Type I SS are often not appropriate. Check the SAS code with this lecture to illustrate the computation of these sums of squares.