CHAPTER OVERVIEW

14: ANOVA Designs, Multiple Factors

Introduction

In our previous discussions about t-tests and ANOVA we focused on procedures with one dependent (response) variable and a single independent (predictor) factor variable that may cause variation in the response variable. In this chapter we extend our discussions about the **general linear model** by

1. Reviewing the one-way ANOVA, and providing a few examples of the one-way design.
2. Reviewing and setting the stage for adding a second independent variable to the model.

Additional one-way ANOVA examples

1. In a plants, we may have a response variable like height and one factor variable (location: sun vs. shade) thought to influence plant height (e.g. Aphalo et al 1999).
2. Pulmonary macrophage phagocytosis behavior (response variable) after exposure of toads to clean air or ozone (factor with 2 levels) (Dohm et al. 2005).
3. Monitor weight change on subjects after 6 weeks eating different diet (DASH, control) (Elmer et al. 2006).

All three of the examples are based on the same **statistical model** which may be written as:

$$Y_{ik} = \mu + A_i + \epsilon_{ik}$$

where μ is the grand mean, Y is the response variable and A is the independent variable, or factor, with $k = 1, 2, \ldots K$ levels, groups, or treatments. The total number of experimental units (e.g., subjects) is given by $i = 1, 2, 3, ldots n$. Note that in the first and third examples, because there were only two groups (example 1: k = location, shade; example 3: k = DASH, control). Note that this problem could have been evaluated as an independent sample t-test. For the second example, there were three groups so k = clean air, first ozone level, second ozone level).

Two-way ANOVA with replication

Biology experiments are typically more complicated than a single t-test or one-way ANOVA design can handle; rarely would we conduct an experiment that reflects only one source of variation.

For example, while diet has a profound effect on weight, clearly, activity levels are also important. At a minimum, when considering a weight loss program, we would want to control or monitor activity of the subjects. This is a two-factor model, and the **main effects**, the two factors, were diet (factor A) and activity, (factor B). Both are expected to affect weight loss, and, perhaps, they may do so in complicated ways — an **interaction** (e.g., on DASH diet, weight loss is accelerated when subjects exercise regularly).

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \epsilon_{ijk}$$

The subject of this chapter is the introduction to **two-way ANOVA** designs. In fact, to many, ANOVA design is practically synonymous to a statistician when they think about experimental design (Lindman 1992; Quinn and Keough 2002). As noted by Quinn and Keough (2002) in the preface to their book, "… many biological hypotheses, even deceptively simple ones, are matched by complex statistical models" (p. xv). Once you start adding factor variables there becomes a number of ways in which the groups and experimental units can be distributed, and thus impact the inferences one can make from the ANOVA results. The first statistical model we introduced was the one-way ANOVA. Next, we begin the two-way ANOVA with the **crossed, balanced, fully replicated design**. Along the way we introduce model symbols to help us communicate the design structure and implications of the statistical models.