# Chapter 14

# Analysis of Variance (Two-Way)

Two-way ANOVA examines how two different factors, such as different experimental treatments, affect the means of the different groups. For example, we might be interested in how different baits, as well as trap color, affect the number of insects caught in the traps. If we conducted an experiment where traps were deployed with different combinations of bait and trap color, this would be a **two-way factorial design**, where the term factorial implies all possible combinations of the two factors. If there were three different baits (A, B, and C) and two trap colors (black, white), a factorial design implies there would be six different treatment combinations in the experiment (A-black, A-white, B-black, B-white, C-black, C-white). There would be one or more traps deployed with each treatment combination. It is customary to call one of the factors in a two-way design 'Factor A', while the other is 'Factor B'.

Similar to one-way ANOVA designs, the factors in two-way ANOVA can be either fixed or random. In the insect trapping experiment discussed above, both bait and trap color would be fixed effects because they were selected by the investigator. There are then $F$ tests for each factor in the design, and potentially a test for the **interaction** of the two factors. **An interaction between two factors implies there is a joint effect of the two factors beyond that predicted by each factor operating additively.** For example, insects might be strongly attracted to A-black traps, more than would be predicted by the bait and trap color effects observed in the rest of the treatments. We will focus some effort on the analysis of this design

because it is one of the more common ones.

There are other possible two-way designs, including one fixed and one random effect, or more rarely both effects are random. We will examine a popular design where one factor is fixed and the other random, called a **randomized block design**. There is an $F$ test for the fixed effect in this design, and this test is often the primary goal of the analysis. With respect to the random effects, it is common to simply estimate the variance components associated with these effects and not conduct any tests, although these are still available. This design is ubiquitous in field studies because it helps control for certain forms of spatial or temporal heterogeneity in the observations, permitting a more powerful test of any treatment or group effects.

What do the data look like for a two-way ANOVA design? We will first examine a simplified data set from a trapping study of the bark beetle predator *T. dubius* (Reeve et al. 2009). These predators feed on bark beetles which attack and kill pine trees, and are attracted to the pheromones of the bark beetles as well as odors emitted by damaged pines. Visual cues may also play a role in their behavior, in particular the dark vertical silhouette provided by the bole of the tree. Three different baits were used: frontalin + turpentine (FRT), ipsdienol + turpentine (IDT), and ipsenol + turpentine (IST). Frontalin, ipsdienol, and ipsenol are bark beetle pheromones, while turpentine contains volatiles similar to those in pine resin. The traps were also painted two different colors, black vs. white, to manipulate their appearance to the predators. Thus, there were a total of six treatments (three baits, two colors) in the design. The different treatments were randomly assigned to trapping locations along transects in a pine forest, with four replicates per treatment. The number of predators caught in each trap were counted after several weeks of trapping (Table 14.1). The fourth column in the table shows the values after applying a log transformation, which is commonly used with count data (see Chapter 15).

We will use the notation $Y_{ijk}$ to reference the observations in two-way ANOVA designs. The $i$ subscript refers to the group or treatment within Factor A (bait), $j$ the group or treatment within Factor B (trap color), while $k$ refers to the observation within the treatment. For example, $Y_{123}$ refers to the third observation in the FRT bait - W color treatment, which is 0.903.

We will also examine data from an experiment that examined how nutrient and water availability, as well as resource heterogeneity in space or time, affect biomass production in grassland plants (Maestre & Reynolds

2007). Plants from a grassland community were seeded in small containers in the greenhouse, with the treatments consisting of different levels of nitrogen and watering. There were three nitrogen and three watering levels in the experiment, for a total of nine treatments, with four replicate containers per treatment. The experiment also included treatments were the nitrogen was heterogeneously distributed in the container and watering was pulsed in time, but we will defer analysis of these other factors to Chapter 19. The total biomass of the plants was then determined after 100 d of growth (Table 14.2).

The data sets presented in this chapter are balanced designs with the same number of replicates per group, because this simplifies the formulas. They can be extended to unbalanced designs, but we will let SAS handle the details of the calculations in this case. We will later see how unbalanced data sets can influence the tests in two-way ANOVA.

Table 14.1: Example 1 - Effect of bait and trap color on catches of *T. dubius*, a bark beetle predator (Reeve et al. 2009). The baits used were frontalin + turpentine (FRT), ipsdienol + turpentine (IDT), and ipsenol + turpentine (IST), and the traps were painted either black (B) or white (W). Also shown are the means for each treatment group ($\bar{Y}_{ij.}$) and preliminary calculations to find $SS_{within}$

| Bait | Color | *T. dubius* | $Y_{ijk} = \log_{10}(T.dubius + 1)$ | $i$ | $j$ | $k$ | $\bar{Y}_{ij.}$ | $(Y_{ijk} - \bar{Y}_{ij.})^2$ |
|------|-------|-------------|-------------------------------------|-----|-----|-----|-----------------|-------------------------------|
| FRT | B | 18 | 1.279 | 1 | 1 | 1 | 1.150 | $1.664 \times 10^{-2}$ |
| FRT | B | 12 | 1.114 | 1 | 1 | 2 | | $1.296 \times 10^{-3}$ |
| FRT | B | 22 | 1.362 | 1 | 1 | 3 | | $4.494 \times 10^{-2}$ |
| FRT | B | 6 | 0.845 | 1 | 1 | 4 | | $9.303 \times 10^{-2}$ |
| FRT | W | 12 | 1.114 | 1 | 2 | 1 | 0.980 | $1.796 \times 10^{-2}$ |
| FRT | W | 15 | 1.204 | 1 | 2 | 2 | | $5.018 \times 10^{-2}$ |
| FRT | W | 7 | 0.903 | 1 | 2 | 3 | | $5.929 \times 10^{-3}$ |
| FRT | W | 4 | 0.699 | 1 | 2 | 4 | | $7.896 \times 10^{-2}$ |
| IDT | B | 0 | 0.000 | 2 | 1 | 1 | 0.369 | $1.363 \times 10^{-1}$ |
| IDT | B | 2 | 0.477 | 2 | 1 | 2 | | $1.161 \times 10^{-2}$ |
| IDT | B | 1 | 0.301 | 2 | 1 | 3 | | $4.658 \times 10^{-3}$ |
| IDT | B | 4 | 0.699 | 2 | 1 | 4 | | $1.087 \times 10^{-1}$ |
| IDT | W | 2 | 0.477 | 2 | 2 | 1 | 0.314 | $2.665 \times 10^{-2}$ |
| IDT | W | 1 | 0.301 | 2 | 2 | 2 | | $1.626 \times 10^{-4}$ |
| IDT | W | 2 | 0.477 | 2 | 2 | 3 | | $2.665 \times 10^{-2}$ |
| IDT | W | 0 | 0.000 | 2 | 2 | 4 | | $9.844 \times 10^{-2}$ |

| Bait | Color | $T.\ dubius$ | $Y_{ijk} = \log_{10}(T.dubius + 1)$ | $i$ | $j$ | $k$ | $\bar{Y}_{ij\cdot}$ | $(Y_{ijk} - \bar{Y}_{ij\cdot})^2$ |
|------|-------|-----|-------|---|---|---|-------|-----------------------|
| IST | B | 2 | 0.477 | 3 | 1 | 1 | 0.725 | $6.126 \times 10^{-2}$ |
| IST | B | 2 | 0.477 | 3 | 1 | 2 | | $6.126 \times 10^{-2}$ |
| IST | B | 10 | 1.041 | 3 | 1 | 3 | | $1.002 \times 10^{-1}$ |
| IST | B | 7 | 0.903 | 3 | 1 | 4 | | $3.186 \times 10^{-2}$ |
| IST | W | 1 | 0.301 | 3 | 2 | 1 | 0.719 | $1.745 \times 10^{-1}$ |
| IST | W | 4 | 0.699 | 3 | 2 | 2 | | $3.901 \times 10^{-4}$ |
| IST | W | 14 | 1.176 | 3 | 2 | 3 | | $2.091 \times 10^{-1}$ |
| IST | W | 4 | 0.699 | 3 | 2 | 4 | | $3.901 \times 10^{-4}$ |

Table 14.2: Example 2 - Effect of nutrient and water availability on the total biomass of grassland plants grown in microcosms (Maestre & Reynolds 2007).

| N (mg) | Water (ml/week) | $Y_{ijk}$ = Biomass | $i$ | $j$ | $k$ |
|--------|-----------------|---------------------|-----|-----|-----|
| 40 | 125 | 4.372 | 1 | 1 | 1 |
| 40 | 125 | 4.482 | 1 | 1 | 2 |
| 40 | 125 | 4.221 | 1 | 1 | 3 |
| 40 | 125 | 3.977 | 1 | 1 | 4 |
| 40 | 250 | 7.400 | 1 | 2 | 1 |
| 40 | 250 | 8.027 | 1 | 2 | 2 |
| 40 | 250 | 7.883 | 1 | 2 | 3 |
| 40 | 250 | 7.769 | 1 | 2 | 4 |
| 40 | 375 | 7.226 | 1 | 3 | 1 |
| 40 | 375 | 8.126 | 1 | 3 | 2 |
| 40 | 375 | 6.840 | 1 | 3 | 3 |
| 40 | 375 | 7.901 | 1 | 3 | 4 |
| 80 | 125 | 5.140 | 2 | 1 | 1 |
| 80 | 125 | 3.913 | 2 | 1 | 2 |
| 80 | 125 | 4.669 | 2 | 1 | 3 |
| 80 | 125 | 4.306 | 2 | 1 | 4 |
| 80 | 250 | 9.099 | 2 | 2 | 1 |
| 80 | 250 | 9.711 | 2 | 2 | 2 |
| 80 | 250 | 9.123 | 2 | 2 | 3 |
| 80 | 250 | 9.709 | 2 | 2 | 4 |
| 80 | 375 | 10.701 | 2 | 3 | 1 |
| 80 | 375 | 11.552 | 2 | 3 | 2 |
| 80 | 375 | 11.356 | 2 | 3 | 3 |
| 80 | 375 | 9.759 | 2 | 3 | 4 |

| N (mg) | Water (ml) | $Y_{ijk} =$ Biomass | $i$ | $j$ | $k$ |
|--------|-----------|---------------------|-----|-----|-----|
| 120 | 125 | 5.021 | 3 | 1 | 1 |
| 120 | 125 | 4.970 | 3 | 1 | 2 |
| 120 | 125 | 5.055 | 3 | 1 | 3 |
| 120 | 125 | 4.862 | 3 | 1 | 4 |
| 120 | 250 | 9.029 | 3 | 2 | 1 |
| 120 | 250 | 10.791 | 3 | 2 | 2 |
| 120 | 250 | 9.115 | 3 | 2 | 3 |
| 120 | 250 | 10.319 | 3 | 2 | 4 |
| 120 | 375 | 12.189 | 3 | 3 | 1 |
| 120 | 375 | 14.381 | 3 | 3 | 2 |
| 120 | 375 | 13.153 | 3 | 3 | 3 |
| 120 | 375 | 14.066 | 3 | 3 | 4 |

## 14.1   Random assignment of treatments

**A essential step in executing ANOVA designs is the random assignment of treatments to experimental units.** For instance, in the Example 2 experiment we would want to randomly assign nitrogen and watering levels to the microcosms. This avoids any bias on the part of the experimenter in assigning the treatments to the containers, and also ensures that the replicates for each treatment are spread and intermingled throughout the greenhouse. What could happen if the treatments are not randomly assigned? Suppose that all the replicates for a given treatment in Example 2 are placed next to each other in the greenhouse, perhaps because this is convenient when applying the treatments. If a particular location happens to be warmer or receive more sunlight than another location, then the plants may be larger in that location and so bias the results of the experiment. We may falsely conclude a particular treatment has an effect on biomass because of this location effect. The random assignment of treatments avoids biases of this sort and also ensures independence of the observations, a basic assumption of most statistical models (Hurlbert 1984; Potvin 1993). Experiments with this feature are also known as **completely randomized designs**. We will illustrate the random assignment of treatments using a SAS program below.

### 14.1.1   Random assignment of treatments - SAS Demo

The program below shows one way of randomly assigning treatments to containers for the Example 2 experiment. We first input the different treatment combinations using a `data` step, with one line in the data set for each replicate. The `data` step also assigns a random number to each observation. The program uses a uniform random variable generated by the `ranuni` function, but any continuous random variable would work. We then use `proc sort` to sort the observations in ascending order by this random variable, thereby randomly shuffling the treatments (see Fig. 14.2). We would then assign to the first container the first treatment combination in the shuffled observations, the second container the second treatment combination, and so forth.

```
* Rand_treatments.sas;
title "Random assignment of Example 2 treatments";
data treat;
    input nitrogen water;
    * Generate a uniform random variable;
    u = ranuni(0);
    datalines;
 40  125
 40  125
 40  125
 40  125
 40  250
 40  250
 40  250
 40  250

etc.

;
run;
title2 "Original order of treatments";
proc print data=treat;
run;
* Sort treatments by value of u;
proc sort out=shuffled data=treat;
    by u;
run;
title2 "Randomly shuffled treatments";
proc print data=shuffled;
run;
```

```
quit;
```

**Random assignment of Example 2 treatments**
**Original order of treatments**

| Obs | nitrogen | water | u |
|---|---|---|---|
| 1 | 40 | 125 | 0.73857 |
| 2 | 40 | 125 | 0.74237 |
| 3 | 40 | 125 | 0.78767 |
| 4 | 40 | 125 | 0.91983 |
| 5 | 40 | 250 | 0.08052 |
| 6 | 40 | 250 | 0.35282 |
| 7 | 40 | 250 | 0.11267 |
| 8 | 40 | 250 | 0.14701 |
| 9 | 40 | 375 | 0.52189 |
| 10 | 40 | 375 | 0.87401 |

etc.

Figure 14.1: `rand_treatments.sas` - `proc print`

**Random assignment of Example 2 treatments**
**Randomly shuffled treatments**

| Obs | nitrogen | water | u |
|---|---|---|---|
| 1 | 80 | 375 | 0.02958 |
| 2 | 120 | 250 | 0.03010 |
| 3 | 40 | 250 | 0.08052 |
| 4 | 40 | 250 | 0.11267 |
| 5 | 40 | 250 | 0.14701 |
| 6 | 120 | 375 | 0.17509 |
| 7 | 80 | 250 | 0.19177 |
| 8 | 120 | 125 | 0.19541 |
| 9 | 40 | 375 | 0.20295 |
| 10 | 80 | 125 | 0.25478 |

etc.

Figure 14.2: `rand_treatments.sas` - `proc print`

## 14.2 Two-way fixed effects model

Suppose that we want to model the observations in studies like Example 1 or 2, where there are two factors that are manipulated and are fixed effects. Let Factor A be one treatment (such as bait type) while Factor B is the other treatment (trap color). Let the symbol $Y_{ijk}$ stand for the *kth* observation ($k = 1, 2, \ldots, n$) in the *ith* Factor A treatment and *jth* Factor B treatment. For example, with the Example 1 data set we have $Y_{111} = 1.279$ while $Y_{222} = 0.301$ (see Table 14.1). One commonly used model for such a design (Searle 1971) is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}. \tag{14.1}$$

Here $\mu$ is the grand mean of the observations, while $\alpha_i$ is the deviation from $\mu$ caused by the *ith* treatment in Factor A, while $\beta_j$ is the deviation caused by the *jth* treatment in Factor B. These terms are called the **main effects** in the model. The term $(\alpha\beta)_{ij}$ represents an **interaction** between Factors A and B, implying a shift in the mean for a particular treatment combination beyond the effects of Factor A and B. An interaction between two factors A and B is often symbolized as 'A $\times$ B.' It is also considered a fixed effect when both A and B are fixed effects. The $\epsilon_{ijk}$ term represents random departures from the mean value predicted by the main effects and interaction due to natural variability among the observations, and are also assumed to be independent. The model also assumes that $\sum \alpha_i = 0$, $\sum \beta_j = 0$, and $\sum (\alpha\beta)_{ij} = 0$, but this does not affect its generality. The same model can also be used to describe the observations for studies where there are $a$ groups or levels for Factor A, and $b$ for Factor B, with any number of replicates ($n$) per treatment combination, as well as unbalanced designs with different numbers of replicates.

It follows for the *ith* level of Factor A and *jth* of Factor B that $E[Y_{ijk}] = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ and $Var[Y_{ijk}] = \sigma^2$, using the rules for expected values and variances. Thus, for the *ith* and *jth* level we have $Y_{ijk} \sim N(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}, \sigma^2)$. We can illustrate how the different parameters work in this model by plotting the distribution of the data for different parameter values. The behavior of the model is described for four different scenarios below. We will model an experiment similar to Example 1, where there are three levels for Factor A ($a = 3$) and two for Factor B ($b = 2$).

### 14.2.1    Factor A effect

Suppose that Factor A has a strong effect on $Y_{ijk}$, but there is only a minimal effect of Factor B and no interaction between the two factors. To make things concrete, let $\mu = 1.5$, $\alpha_1 = 0.5, \alpha_2 = 0, \alpha_3 = -0.5, \beta_1 = 0.1, \beta_2 = -0.1$, $(\alpha\beta)_{ij} = 0$ for all $i$ and $j$, and $\sigma^2 = 0.05$. Figure 14.3 shows the distribution of the observations in each treatment group. Note that the mean for treatment 1 under Factor A is shifted upward from $\mu$ while treatment 3 is shifted downward, for both levels of Factor B. The distribution for each treatment combination has the same variance, namely $\sigma^2 = 0.05$.

### 14.2.2    Factor B effect

Suppose the reverse situation is now true, with Factor B having a strong effect on $Y_{ijk}$ while Factor A has a minimal effect, again with no interaction. This could be modeled using $\alpha_1 = 0.1, \alpha_2 = 0, \alpha_3 = -0.1, \beta_1 = 0.5, \beta_2 = -0.5$, and $(\alpha\beta)_{ij} = 0$ for all $i$ and $j$. Figure 14.4 shows the pattern that results. Note that the mean for treatment 1 under Factor B is shifted upward from $\mu$, while treatment 2 is shifted downward, for all three levels of Factor A.

### 14.2.3    Factor A and B effect

If both factors have an effect on $Y_{ijk}$, we would expect to see a combination of the previous patterns, with the treatment groups shifted away from each other (Fig. 14.5). This figure uses $\alpha_1 = 0.5, \alpha_2 = 0, \alpha_3 = -0.5, \beta_1 = 0.3, \beta_2 = -0.3$, and $(\alpha\beta)_{ij} = 0$ for all $i$ and $j$.

### 14.2.4    Interaction effect

We now examine how an A $\times$ B interaction influences the model. Suppose that $\alpha_1 = 0.5, \alpha_2 = 0, \alpha_3 = -0.5, \beta_1 = 0.3$ and $\beta_2 = -0.3$ as in the previous figure, but now $(\alpha\beta)_{11} = 0.2, (\alpha\beta)_{12} = -0.2, (\alpha\beta)_{21} = 0, (\alpha\beta)_{22} = 0, (\alpha\beta)_{31} = -0.2$, and $(\alpha\beta)_{32} = 0.2$. We see that Factor B has a substantial effect under treatment 1 for Factor A, and smaller effect under treatment 2, and almost no effect under treatment 3 (Fig. 14.6). Note that the distributions under the different treatment combinations no longer move in parallel as in Fig. 14.5. This pattern is diagnostic of an interaction in the analysis of

real data. We will later examine a data set where there is strong interaction between the two factors.

The objective in two-way ANOVA is to test whether Factor A, B, or both have an effect on the group means, and whether there is interaction between the two factors. For Factor A this amounts to testing $H_0$ : all $\alpha_i = 0$, while for Factor B we would test $H_0$ : all $\beta_j = 0$. For interaction between the two factors, we would test $H_0$ : all $(\alpha\beta)_{ij} = 0$. The corresponding alternative hypotheses are $H_1$ : some $\alpha_i \neq 0$, $H_1$ : some $\beta_j \neq 0$, and $H_1$ : some $(\alpha\beta)_{ij} \neq 0$. We will discuss how these null hypotheses are tested in the next section.
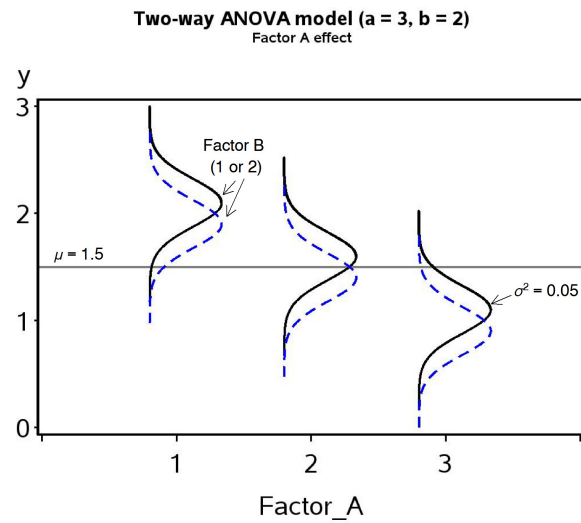
Figure 14.3: Fixed effects model for two-way ANOVA showing a Factor A effect.
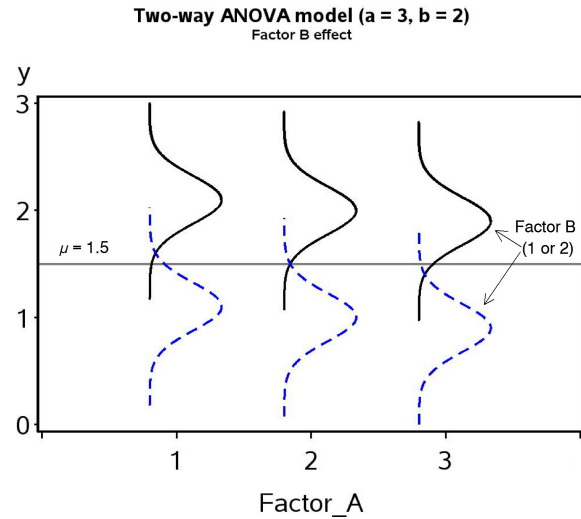


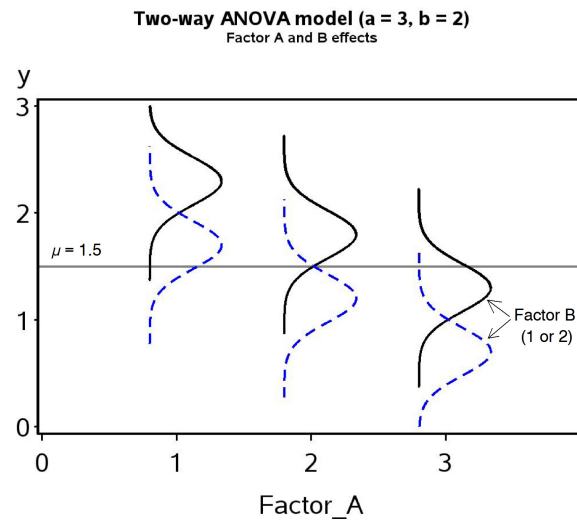Figure 14.4: Fixed effects model for two-way ANOVA showing a Factor B effect.

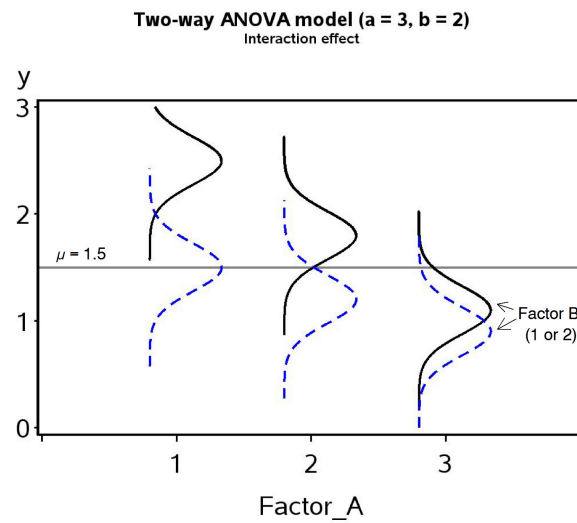Figure 14.5: Fixed effects model for two-way ANOVA showing both Factor A and B effects.



Figure 14.6: Fixed effects model for two-way ANOVA showing an A × B interaction between the two factors.

## 14.3     Hypothesis testing for two-way ANOVA

We now develop statistical tests for each of the null hypotheses listed above. All work in a similar fashion to the $F$ test for one-way ANOVA. For Factor A and B in the model, as well as the interaction term, there is a corresponding sum of squares and mean square term. There is also an overall sum of squares and mean square within groups. These quantities are used to construct three different $F$ tests, one for Factor A, Factor B, and the A $\times$ B interaction. These three tests are also examples of likelihood ratio tests, in which the fit is compared between the null and alternative models (Searle 1971). We will illustrate the calculations for these tests using the Example 1 data set, with Factor A being bait while Factor B is trap color.

### 14.3.1     Sum of squares and mean squares

We begin by calculating the group means for each treatment combination. For the Example 1 data, this amounts to calculating a group mean for each combination of bait and trap color. These group means are shown in Table 14.1 and labeled as $\bar{Y}_{ij\cdot}$. Here the '$\cdot$' notation implies the mean was calculated using all the observations in that group $(k = 1, 2, \ldots, n)$. A grand mean can then be calculated as the mean of these group means, or equivalently by summing all the observations and dividing by their total number. We label this grand mean as $\bar{\bar{Y}}$. It can be generally calculated using the formula

$$\bar{\bar{Y}} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} \bar{Y}_{ij\cdot}}{ab}. \tag{14.2}$$

For the Example 1 data set, we have

$$\bar{\bar{Y}} = \frac{1.150 + 0.980 + 0.369 + 0.314 + 0.725 + 0.719}{6} = 0.709. \tag{14.3}$$

We next calculate a mean corresponding to each level of Factor A by averaging across the levels of Factor B, which we denote as $\bar{\bar{Y}}_{i\cdot\cdot}$. It can be calculated using the formula

$$\bar{\bar{Y}}_{i\cdot\cdot} = \frac{\sum_{j=1}^{b} \bar{Y}_{ij\cdot}}{b}. \tag{14.4}$$

For the Example 1 data set, we have

$$\bar{\bar{Y}}_{1\cdot\cdot} = \frac{1.150 + 0.980}{2} = 1.065, \tag{14.5}$$

$$\bar{\bar{Y}}_{2..} = \frac{0.369 + 0.314}{2} = 0.342, \tag{14.6}$$

and

$$\bar{\bar{Y}}_{3..} = \frac{0.725 + 0.719}{2} = 0.722. \tag{14.7}$$

The difference $\bar{\bar{Y}}_{i..} - \bar{\bar{\bar{Y}}}$ is a measure of the shift generated by Factor A in the observations, as well as an estimate of $\alpha_i$ for each level of Factor A. We can obtain a single measure of these shifts by squaring and summing them across all groups to obtain a sum of squares for Factor A, or $SS_A$. It can be calculated using the general formula

$$SS_A = nb \sum_{i=1}^{a} (\bar{\bar{Y}}_{i..} - \bar{\bar{\bar{Y}}})^2. \tag{14.8}$$

$SS_A$ has $a-1$ degrees of freedom. We can calculate a mean square for Factor A using the formula

$$MS_A = \frac{SS_A}{a-1}. \tag{14.9}$$

Note the factor $nb$ in the expression for $SS_A$, which scales $MS_A$ so that it estimates $\sigma^2$ if $H_0$ : all $\alpha_i = 0$ is true (no Factor A effect). If $H_1$ is true, implying some $\alpha_i \neq 0$, then $MS_A$ will become larger. For the Example 1 data, we have

$$SS_A = 4(2) \left[ (1.065 - 0.709)^2 + (0.342 - 0.709)^2 + (0.722 - 0.709)^2 \right] \tag{14.10}$$

$$= 8 \left[ 1.265 \times 10^{-1} + 1.353 \times 10^{-1} + 1.501 \times 10^{-4} \right] = 2.096. \tag{14.11}$$

and

$$MS_A = \frac{2.096}{3-1} = 1.048. \tag{14.12}$$

We can similarly calculate a mean corresponding to each level of Factor B, averaging across levels of Factor A. The general formula for these means is

$$\bar{\bar{Y}}_{.j.} = \frac{\sum_{i=1}^{a} \bar{Y}_{ij.}}{a}. \tag{14.13}$$

For the Example 1 data set, we have

$$\bar{\bar{Y}}_{.1.} = \frac{1.150 + 0.369 + 0.725}{3} = 0.748 \tag{14.14}$$

and

$$\bar{\bar{Y}}_{\cdot 2 \cdot} = \frac{0.980 + 0.314 + 0.719}{3} = 0.671. \tag{14.15}$$

The difference $\bar{\bar{Y}}_{\cdot j \cdot} - \bar{\bar{\bar{Y}}}$ is a measure of the shift generated by Factor B in the observations, as well as an estimate of $\beta_j$ for each level of Factor B. Squaring and summing them across all groups, we obtain a sum of squares for Factor B, or $SS_B$. It can be calculated using the general formula

$$SS_B = na \sum_{j=1}^{b} (\bar{\bar{Y}}_{\cdot j \cdot} - \bar{\bar{\bar{Y}}})^2. \tag{14.16}$$

$SS_B$ has $b - 1$ degrees of freedom. We can then calculate a mean square for Factor B using the formula

$$MS_B = \frac{SS_B}{b - 1}. \tag{14.17}$$

For the Example 1 data, we have

$$SS_B = 4(3) \left[ (0.748 - 0.709)^2 + (0.671 - 0.709)^2 \right] \tag{14.18}$$

$$= 12 \left[ 1.485 \times 10^{-3} + 1.485 \times 10^{-3} \right] = 3.565 \times 10^{-2} \tag{14.19}$$

and

$$MS_B = \frac{3.565 \times 10^{-2}}{2 - 1} = 3.565 \times 10^{-2}. \tag{14.20}$$

We can also calculate a sum of squares and mean square to test for the A $\times$ B interaction. The sum of squares for interaction, $SS_{AB}$, is calculated in general using the formula

$$SS_{AB} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{Y}_{ij\cdot} - \bar{\bar{Y}}_{i\cdot\cdot} - \bar{\bar{Y}}_{\cdot j\cdot} + \bar{\bar{\bar{Y}}})^2. \tag{14.21}$$

The terms within this expression estimate $(\alpha\beta)_{ij}$, and are measures of the difference between the means for each treatment combination and the values predicted by the model without any interaction. $SS_{AB}$ has $(a - 1)(b - 1)$ degrees of freedom. Its associated mean square is defined by the formula

$$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}. \tag{14.22}$$

For the Example 1 data, we have

$$
\begin{aligned}
SS_{AB} &= 4[(1.150 - 1.065 - 0.748 + 0.709)^2 & (14.23)\\
&\quad +(0.980 - 1.065 - 0.671 + 0.709)^2 & (14.24)\\
&\quad +(0.369 - 0.342 - 0.748 + 0.709)^2 & (14.25)\\
&\quad +(0.314 - 0.342 - 0.671 + 0.709)^2 & (14.26)\\
&\quad +(0.725 - 0.722 - 0.748 + 0.709)^2 & (14.27)\\
&\quad +(0.719 - 0.722 - 0.671 + 0.709)^2 & (14.28)\\
&= 5[2.111 \times 10^{-3} + \cdots + 2.836 \times 10^{-2}] = 2.836 \times 10^{-2}. & (14.29)
\end{aligned}
$$

and

$$ MS_{AB} = \frac{2.836 \times 10^{-2}}{(3-1)(2-1)} = 1.418 \times 10^{-2}. \tag{14.30} $$

These sum of squares and mean squares measure how Factor A, B, and the A × B interaction influence the means of each treatment combination. What about variability within each group? We can calculate $SS_{within}$ using the general formula

$$ SS_{within} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \bar{Y}_{ij\cdot})^2 \tag{14.31} $$

which has $ab(n-1)$ degrees of freedom. The associated mean square is calculated as

$$ MS_{within} = \frac{SS_{within}}{ab(n-1)}. \tag{14.32} $$

The last column of Table 14.1 shows the preliminary calculations for $SS_{within}$. Adding this column across all the treatment groups yields

$$ SS_{within} = 1.644 \times 10^{-2} + \cdots + 3.901 \times 10^{-4} = 1.361 \tag{14.33} $$

and

$$ MS_{within} = \frac{1.361}{(3)(2)(4-1)} = 7.561 \times 10^{-2}. \tag{14.34} $$

There is one more sum of squares that is often calculated in two-way ANOVA, the total sum of squares. It is defined as

$$ SS_{total} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \bar{\bar{Y}})^2. \tag{14.35} $$

It measures the variability of the observations around the grand mean of the data ($\bar{\bar{\bar{Y}}}$) and has $abn - 1$ degrees of freedom. An interesting feature of the sum of squares is that they add to the total sum of squares when the design is balanced, as do the degrees of freedom. In particular, we have

$$SS_A + SS_B + SS_{AB} + SS_{within} = SS_{total} \qquad (14.36)$$

and

$$(a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1) = abn - 1. \qquad (14.37)$$

Thus, the sum of squares and degrees of freedom can be partitioned into components corresponding to every source of variation in the study. For Example 1, we have $SS_{total} = 3.521$ with $3(2)(4) - 1 = 23$ degrees of freedom.

## 14.3.2   ANOVA tables and tests

We can organize the different sum of squares and mean squares into an ANOVA table. It lists the different sources of variation in the data (Factor A, B, A × B interaction, within groups, and total) and their degrees of freedom. Table 14.3 shows the general layout of such a table for two-way ANOVA designs.

Also shown in the table are $F$ statistics used test to whether Factor A, Factor B, and their interaction have an effect on the observations. The numerator of the test statistic is the mean square for each factor ($MS_A$, $MS_B$, or $MS_{AB}$), while the denominator is always $MS_{within}$. Thus, we use $F_s = MS_A/MS_{within}$ to test for the effect of Factor A. Under $H_0$ : all $\alpha_i = 0$ this statistic has an $F$ distribution with $df_1 = a - 1$ and $df_2 = ab(n - 1)$. Similarly, we use $F_s = MS_B/MS_{within}$ to test for an effect of Factor B. Under $H_0$ : all $\beta_j = 0$ it has an $F$ distribution with $df_1 = b - 1$ and $df_2 = ab(n - 1)$. Finally, we use $F_s = MS_{AB}/MS_{within}$ to test for an interaction between A and B. Under $H_0$ : all $(\alpha\beta)_{ij} = 0$ it has an $F$ distribution with $df_1 = (a - 1)(b - 1)$ and $df_2 = ab(n - 1)$.

All these tests are examples of likelihood ratio tests. For example, consider the test for the A × B interaction. To construct the likelihood ratio test for the interaction, we first find the maximum likelihood estimates of various parameters under $H_1$ vs. $H_0$. Recall that the observations in the two-way ANOVA model are described as

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}. \qquad (14.38)$$

where $\mu$ is the grand mean, $\alpha_i$ is the effect of the *ith* level of Factor A, $\beta_j$ is the effect of the *jth* level of Factor B, $(\alpha\beta)_{ij}$ is effect of the interaction, and $\epsilon_{ijk} \sim N(0, \sigma^2)$. This is the statistical model under the alternative hypothesis $H_1$ : some $(\alpha\beta)_{ij} \neq 0$, implying an interaction effect. Under $H_0$ : all $(\alpha\beta)_{ij} = 0$, the model reduces to

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}. \qquad (14.39)$$

We would need to find the maximum likelihood estimates under both $H_1$ and $H_0$, as well as $L_{H_0}$ and $L_{H_1}$, the maximum height of the likelihood function under $H_0$ and $H_1$. We would then use the likelihood ratio test statistic

$$\lambda = \frac{L_{H_0}}{L_{H_1}}. \qquad (14.40)$$

It can be shown that there is a one-to-one correspondence between $-2\ln(\lambda)$ and $F_s$ for the interaction effect, and so the $F$ test is actually a likelihood ratio test (Searle 1971), as are the tests for the other effects. Large values of the test statistic $-2\ln(\lambda)$ or $F_s$ indicate a lower value of the likelihood under $H_0$ relative to $H_1$, and thus a poorer fit of the $H_0$ model.

Table 14.4 shows the results for the Example 1 data set, including the $F$ statistics and $P$ values obtained using Table F. **In examining the test results, it is customary to examine the test for the interaction first, followed by the main effects.** If the interaction is nonsignificant this suggests the two main effects have a simple additive effect on the observations, provided they are significant. If the interaction is significant the interpretation requires more attention. If one or more of the main effects are significant, it suggests the observations are driven by both interaction and main effects. Fig. 14.6 shows a theoretical example where an interaction, Factor A, and Factor B all influence the observations.

For the bait $\times$ trap color interaction, we see that $F_s = 0.19$ with $df_1 = 2$ and $df_2 = 18$, and from Table F find that $P > 0.100$. Thus, the interaction was nonsignificant for these data ($F_{2,18} = 0.19, P > 0.100$). The color effect was also nonsignificant ($F_{1,18} = 0.47, P > 0.100$), but the bait effect was highly significant ($F_{2,18} = 13.86, P < 0.001$). Each bait represents a different bark beetle pheromone, and apparently some baits are more attractive than others for *T. dubius*.

Table 14.3: General ANOVA table for two-way designs with replication, showing formulas for different mean squares and $F$ tests.

| Source | df | Sum of squares | Mean square | $F_s$ |
|---|---|---|---|---|
| Factor A | $a-1$ | $SS_A$ | $MS_A = SS_A/(a-1)$ | $MS_A/MS_{within}$ |
| Factor B | $b-1$ | $SS_B$ | $MS_B = SS_B/(b-1)$ | $MS_B/MS_{within}$ |
| AB interaction | $(a-1)(b-1)$ | $SS_{AB}$ | $MS_{AB} = SS_{AB}/(a-1)(b-1)$ | $MS_{AB}/MS_{within}$ |
| Within | $ab(n-1)$ | $SS_{within}$ | $MS_{within} = SS_{within}/ab(n-1)$ | |
| Total | $abn-1$ | $SS_{total}$ | | |

Table 14.4: ANOVA table for the Example 1 data set, including $P$ values for the tests.

| Source | df | Sum of squares | Mean square | $F_s$ | $P$ |
|---|---|---|---|---|---|
| Bait | 2 | 2.096 | 1.048 | 13.86 | $< 0.001$ |
| Color | 1 | $3.565 \times 10^{-2}$ | $3.565 \times 10^{-2}$ | 0.47 | $> 0.100$ |
| Bait $\times$ Color | 2 | $2.836 \times 10^{-2}$ | $1.418 \times 10^{-2}$ | 0.19 | $> 0.100$ |
| Within | 18 | 1.361 | $7.561 \times 10^{-2}$ | | |
| Total | 23 | 3.521 | | | |

### 14.3.3 Two-way ANOVA for Example 1 - SAS demo

The same calculations for the Example 1 study can be carried out using `proc glm` (SAS Institute Inc. 2018). This procedure is primarily intended for fixed effects ANOVA models, and this study has two fixed effects, bait type and trap color, plus the interaction is also considered a fixed effect.

The first step in the program (see below) is to read in the observations using a `data` step, with one variable denoting the bait treatment (`bait`), another the trap color (`color`), and the third the number of *T. dubius* captured per trap (`Tdubius`). These numbers are then log-transformed using a SAS function to yield the variable `y = log10(Tdubius+1)`. We add one to the observations before taking the log to avoid problems with zeroes.

The data are then plotted using `proc gplot`, with the `bait` treatment on the *x*-axis and separate lines drawn for each color (SAS Institute Inc. 2016). This is accomplished with the command `plot y*bait=color`. The rest of the `gplot` statements control the appearance of the symbols and axes.

The next section of the program conducts the two-way ANOVA using `proc glm`. The `class` statement tells SAS that both `bait` and `color` are used to classify the observations into the six treatment groups. The `model` statement tells SAS the form of the ANOVA model. Recall that the model for fixed effects two-way ANOVA is given by the equation

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \tag{14.41}$$

The $\alpha_i$, $\beta_j$, and $(\alpha\beta)_{ij}$ terms in this model equate directly with the `bait`, `color`, and `bait*color` entries in the `model` statement. The `lsmeans` statement causes `glm` to calculate quantities called least squares means for each level of `bait` and `color`. When the data are balanced these are equivalent to the means for each treatment group, but least squares means have some advantages for unbalanced data and other statistical models. The option `adjust=tukey` requests multiple comparisons among treatments using the Tukey method. This is useful for comparing the different `bait` treatments, but for color there is only one comparison (black vs. white) and in this case would be equivalent to the $F$ test for `color`.

The `proc glm` output provides information similar to that summarized in an ANOVA table (see Fig. 14.9). The degrees of freedom, sum of squares, mean squares, $F$ statistics and $P$ values for the bait, color, and bait $\times$ color interaction are listed near the bottom of the output under `Type III SS`. The degrees of freedom, sum of squares, and mean square for the variation

within groups are labeled as `Error` above this section (this terminology will be explained in Chapter 15). The output labeled `Type I SS` is produced by sequentially fitting the different terms in the model, in the order listed in the `model` statement. Type III sums of squares are more generally useful than Type I for ANOVA designs, although the results are the same when the design is balanced. The output labeled `Model` refers to the combined variation due to bait, color, plus their interaction. The associated $F$ statistic tests whether any or all of these effects influence the observations vs. the null hypothesis that they have no effect. This particular test is not used much with ANOVA designs.

We now examine the results of these tests, beginning with the interaction (Fig. 14.9). We see that the bait $\times$ color interaction was nonsignificant ($F_{2,18} = 0.19, P = 0.8311$). The color effect was also nonsignificant ($F_{1,18} = 0.47, P = 0.5011$), while bait was highly significant ($F_{2,18} = 13.85, P = 0.0002$). Examining the graph and Tukey results (Fig. 14.8, 14.10), we see that predator densities for the FRT and IST treatments were significantly higher than for IDT. Note that the lines connecting the different treatments are roughly parallel, further indicating an absence of interaction. The effect of trap color appears minimal in this study, although trap catches were somewhat higher for black traps.

──────────────── SAS Program ────────────────

```
* Tdubius_bait_color.sas;
title "Two-way ANOVA for T. dubius trapping";
title2 "Data from Reeve et al. (2009)";
data Tdubius;
    input bait $ color $ Tdubius;
    * Apply transformations here;
    y = log10(Tdubius+1);
    datalines;
FRT B   18
FRT B   12
FRT B   22
FRT B   6
FRT W   12
FRT W   15
FRT W   7
FRT W   4

etc.
```

```
;
run;
* Print data set;
proc print data=Tdubius;
run;
* Plot means, standard errors, and observations;
proc gplot data=Tdubius;
    plot y*bait=color / vaxis=axis1 haxis=axis1 legend=legend1;
    symbol1 i=std1mjt v=star height=2 width=3;
    axis1 label=(height=2) value=(height=2) width=3 major=(width=2) minor=none;
    legend1 label=(height=2) value=(height=2);
run;
* Two-way ANOVA with all fixed effects;
proc glm plots=diagnostics data=Tdubius;
    class bait color;
    model y = bait color bait*color;
    lsmeans bait color / adjust=tukey cl lines;
run;
quit;
```

**Two-way ANOVA for T. dubius counts**
**Data from Reeve et al. (2009)**

| Obs | bait | color | Tdubius | y |
|---:|---|---|---:|---:|
| 1 | FRT | B | 18 | 1.27875 |
| 2 | FRT | B | 12 | 1.11394 |
| 3 | FRT | B | 22 | 1.36173 |
| 4 | FRT | B | 6 | 0.84510 |
| 5 | FRT | W | 12 | 1.11394 |
| 6 | FRT | W | 15 | 1.20412 |
| 7 | FRT | W | 7 | 0.90309 |
| 8 | FRT | W | 4 | 0.69897 |
| 9 | IDT | B | 0 | 0.00000 |
| 10 | IDT | B | 2 | 0.47712 |

etc.

Figure 14.7: `Tdubius_bait_color.sas - proc print`
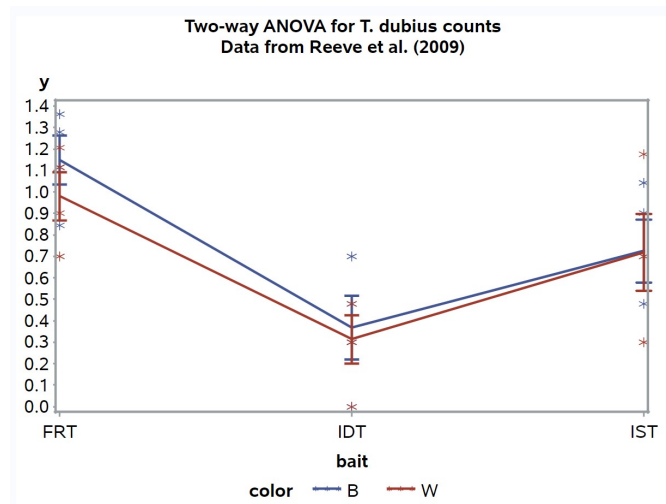


Figure 14.8: `Tdubius_bait_color.sas - proc gplot`

**Two-way ANOVA for T. dubius counts**
**Data from Reeve et al. (2009)**

**The GLM Procedure**

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| bait | 3 | FRT IDT IST |
| color | 2 | B W |

| | |
|---|---|
| Number of Observations Read | 24 |
| Number of Observations Used | 24 |

**Two-way ANOVA for T. dubius counts**
**Data from Reeve et al. (2009)**

**The GLM Procedure**

**Dependent Variable: y**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 5 | 2.15900779 | 0.43180156 | 5.71 | 0.0025 |
| Error | 18 | 1.36120842 | 0.07562269 | | |
| Corrected Total | 23 | 3.52021621 | | | |

| R-Square | Coeff Var | Root MSE | y Mean |
|---|---|---|---|
| 0.613317 | 38.76405 | 0.274996 | 0.709409 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| bait | 2 | 2.09508772 | 1.04754386 | 13.85 | 0.0002 |
| color | 1 | 0.03564427 | 0.03564427 | 0.47 | 0.5011 |
| bait*color | 2 | 0.02827579 | 0.01413790 | 0.19 | 0.8311 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| bait | 2 | 2.09508772 | 1.04754386 | 13.85 | 0.0002 |
| color | 1 | 0.03564427 | 0.03564427 | 0.47 | 0.5011 |
| bait*color | 2 | 0.02827579 | 0.01413790 | 0.19 | 0.8311 |

Figure 14.9: `Tdubius_bait_color.sas` – `proc glm`

**y Tukey Grouping for LS-Means of bait (Alpha = 0.05)**

LS-means covered by the same bar are not significantly different.

| bait | Estimate |
|------|----------|
| FRT | 1.0650 |
| IST | 0.7217 |
| IDT | 0.3415 |

**y Tukey Grouping for LS-Means of color (Alpha = 0.05)**

LS-means covered by the same bar are not significantly different.

| color | Estimate |
|-------|----------|
| B | 0.7479 |
| W | 0.6709 |

Figure 14.10: `Tdubius_bait_color.sas - proc glm`

### 14.3.4 Two-way ANOVA for Example 2 - SAS demo

We next analyze the Example 2 data set using SAS. These data involve the total biomass of grass plants grown in small containers, where the treatments are nitrogen or water availability. The SAS program is similar to the previous example but with different variable names. Ex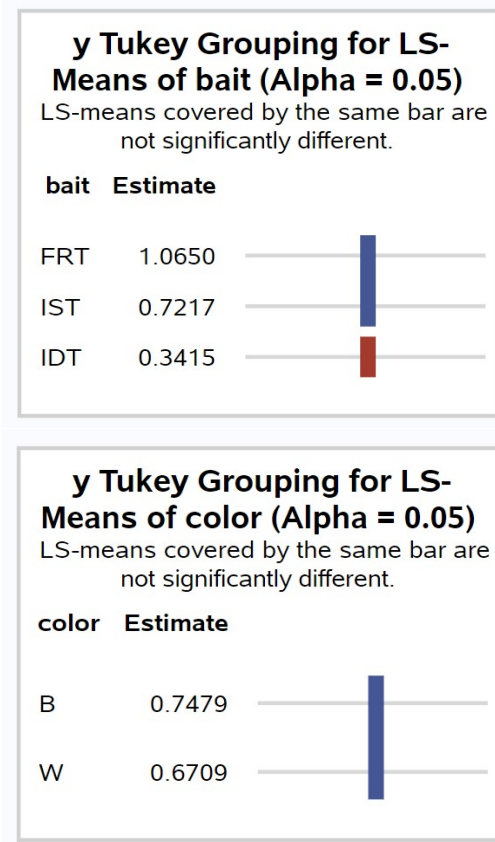amining the output in Fig. 14.13, we see that the nitrogen $\times$ water interaction was highly significant ($F_{4,27} = 11.31, P < 0.0001$). The interaction can be observed in Fig. 14.12, which shows that the lines connecting the treatments are not parallel. Note that the greatest response of biomass to nitrogen occurred at the highest water level, while the response was minimal at the lowest level (Maestre & Reynolds (2007). Thus, low water levels apparently prevent growth even when nitrogen is abundant.

The analysis also found highly significant main effects of nitrogen ($F_{2,27} = 64.28, P < 0.0001$) and water ($F_{2,27} = 456.46, P < 0.0001$) on biomass (Fig. 14.13). There were also significant differences between every nitrogen or water treatment (Fig. 14.14). We can judge the relative strength of these effects by examining Fig. 14.12 as well as their sum of squares values, which are a measure of the amount of variation explained by each effect. They suggest that watering had the most effect on biomass, followed by nitrogen and the nitrogen $\times$ water interaction.

———————————————————————— SAS Program ————————————————————————

```
* Maestre_biomass.sas;
title "Two-way ANOVA for total biomass";
title2 "Data from Maestre and Reynolds (2007)";
data maestre;
    input nitrogen water biomass;
    * Apply transformations here;
    y = log10(biomass);
    datalines;
 40  125    4.372
 40  125    4.482
 40  125    4.221
 40  125    3.977
 40  250    7.400
 40  250    8.027
 40  250    7.883
 40  250    7.769

etc.

;
run;
* Print data set;
proc print data=maestre;
run;
* Plot means, standard errors, and observations;
proc gplot data=maestre;
    plot y*nitrogen=water / vaxis=axis1 haxis=axis1 legend=legend1;
    symbol1 i=std1mjt v=star height=2 width=3;
    axis1 label=(height=2) value=(height=2) width=3 major=(width=2) minor=none;
    legend1 label=(height=2) value=(height=2);
run;
* Two-way ANOVA with all fixed effects;
proc glm plots=diagnostics data=maestre;
    class nitrogen water;
    model y = nitrogen water nitrogen*water;
    lsmeans nitrogen water / adjust=tukey cl lines;
run;
quit;
```

————————————————————————————————————————————————————————————

**Two-way ANOVA for biomass**
**Data from Maestre and Reynolds (2007)**

| Obs | nitrogen | water | biomass | y |
|---|---|---|---|---|
| 1 | 40 | 125 | 4.372 | 0.64068 |
| 2 | 40 | 125 | 4.482 | 0.65147 |
| 3 | 40 | 125 | 4.221 | 0.62542 |
| 4 | 40 | 125 | 3.977 | 0.59956 |
| 5 | 40 | 250 | 7.400 | 0.86923 |
| 6 | 40 | 250 | 8.027 | 0.90455 |
| 7 | 40 | 250 | 7.883 | 0.89669 |
| 8 | 40 | 250 | 7.769 | 0.89037 |
| 9 | 40 | 375 | 7.226 | 0.85890 |
| 10 | 40 | 375 | 8.126 | 0.90988 |

etc.

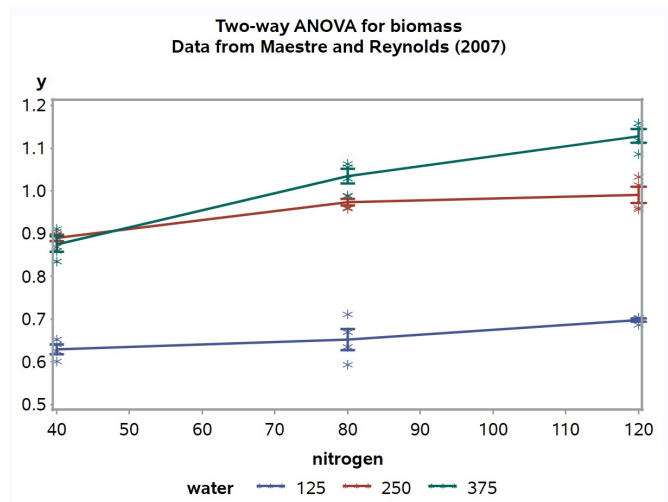Figure 14.11: `Maestre_biomass.sas.sas - proc print`



Figure 14.12: `Maestre_biomass.sas.sas - proc gplot`

## Two-way ANOVA for biomass
## Data from Maestre and Reynolds (2007)

## The GLM Procedure

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| nitrogen | 3 | 40 80 120 |
| water | 3 | 125 250 375 |

| Number of Observations Read | 36 |
|---|---|
| Number of Observations Used | 36 |

## Two-way ANOVA for biomass
## Data from Maestre and Reynolds (2007)

## The GLM Procedure

## Dependent Variable: y

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 8 | 1.01770131 | 0.12721266 | 135.84 | <.0001 |
| Error | 27 | 0.02528594 | 0.00093652 | | |
| Corrected Total | 35 | 1.04298725 | | | |

| R-Square | Coeff Var | Root MSE | y Mean |
|---|---|---|---|
| 0.975756 | 3.499961 | 0.030603 | 0.874368 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| nitrogen | 2 | 0.12039036 | 0.06019518 | 64.28 | <.0001 |
| water | 2 | 0.85496135 | 0.42748068 | 456.46 | <.0001 |
| nitrogen*water | 4 | 0.04234959 | 0.01058740 | 11.31 | <.0001 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| nitrogen | 2 | 0.12039036 | 0.06019518 | 64.28 | <.0001 |
| water | 2 | 0.85496135 | 0.42748068 | 456.46 | <.0001 |
| nitrogen*water | 4 | 0.04234959 | 0.01058740 | 11.31 | <.0001 |

Figure 14.13: `Maestre_biomass.sas - proc glm`

**y Tukey Grouping for LS-Means of nitrogen (Alpha = 0.05)**

LS-means covered by the same bar are not significantly different.

| nitrogen | Estimate | |
|----------|----------|---|
| 120      | 0.9384   | |
| 80       | 0.8864   | |
| 40       | 0.7983   | |

**y Tukey Grouping for LS-Means of water (Alpha = 0.05)**

LS-means covered by the same bar are not significantly different.

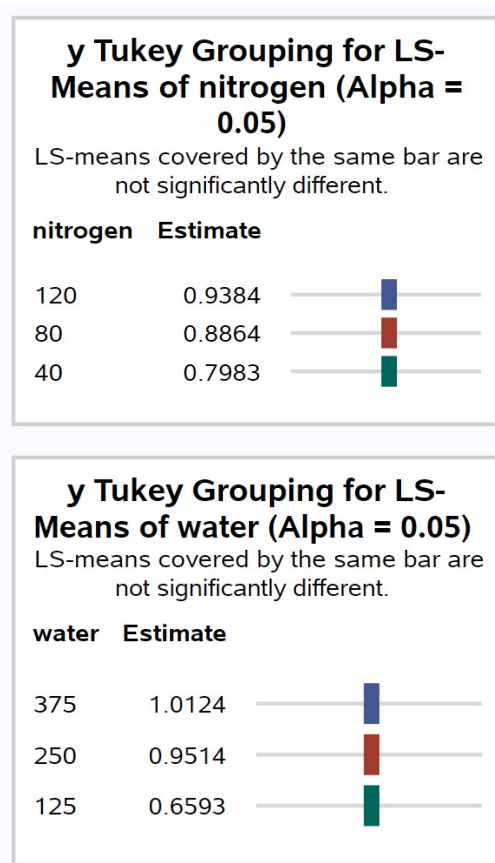| water | Estimate | |
|-------|----------|---|
| 375   | 1.0124   | |
| 250   | 0.9514   | |
| 125   | 0.6593   | |

Figure 14.14: `Maestre_biomass.sas - proc glm`

### 14.3.5    Tests for main effects with interaction

There is disagreement among statisticians on whether tests of the main effects are appropriate when there is significant interaction. Two different procedures have been developed. The SAS one involves fitting models with and without a given main effect, but always including interaction terms, yielding what SAS calls Type III sums of squares and tests (Speed et al. 1978, Shaw & Mitchell-Olds 1993, SAS Institute Inc. 2018). This has the benefit of generating tests for the interaction and main effects in a single pass (see preceding SAS demo). However, there are authors that believe tests of the main effects are questionable in the presence of interaction (e.g., Cox 1984, Winer et al. 1991, Stewart-Oaten 1995). One issue is whether a model with interaction but lacking a main effect is even plausible (Stewart-Oaten 1995). These considerations motivate a different procedure. The first step is to examine the test for interaction using the full two-way ANOVA model. If interaction appears weak or absent, there are two alternate ways of testing the main effects. One is to drop the interaction and rerun the model, examining the main effects in the usual fashion. Another method is to use what SAS calls Type II sums of squares, obtained using the option \ss2 in the `model` statement. The tests based on these sums of squares assume there is no interaction. If the interaction is significant the main effects tests are ignored, although one can still test for Factor A effects at each level of Factor B, or vice versa (Winer et al. 1991). These are called tests of **simple effects**, and can be conducted using the SAS `slice` option for `lsmeans`.

The modified SAS code to implement these procedures is listed below, along with the corresponding output for the Example 2 data set. We see that the nitrogen $\times$ water interaction was highly significant ($F_{4,27} = 11.31, P < 0.0001$), and so we skip the tests of the main effects (Fig. 14.15). Note that the main effects sum of squares are identical to our previous ones using SAS Type III tests, but this would only be true for the special case of balanced designs with equal $n$ for each treatment (see next section for unbalanced designs). The `slice` option is used to test for a nitrogen effect at every level of water, and vice versa (Fig. 14.16). We see that the effect of nitrogen was significant at the lowest water level, while highly significant at the other two levels. It appears the nitrogen effect was smaller at low water levels (see Fig. 14.12). The water effect was highly significant at every level of nitrogen.

―――――――――――――――――――― SAS Program ――――――――――――――――――――

```
* Two-way ANOVA with interaction;
title3 "MODEL WITH INTERACTION - USE THIS OUTPUT IF INTERACTION SIGNIFICANT";
proc glm plots=diagnostics data=maestre;
    class nitrogen water;
    model y = nitrogen water nitrogen*water / ss2;
    lsmeans nitrogen*water / slice=water slice=nitrogen;
run;
* Two-way ANOVA without interaction;
title3 "MODEL WITHOUT INTERACTION - USE THIS OUTPUT IF INTERACTION NS";
proc glm data=maestre;
    class nitrogen water;
    model y = nitrogen water / ss2;
    lsmeans nitrogen water / adjust=tukey cl lines;
run;
```

**Two-way ANOVA for biomass**
**Data from Maestre and Reynolds (2007)**
**MODEL WITH INTERACTION - USE THIS OUTPUT IF INTERACTION SIGNIFICANT**

**The GLM Procedure**

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| nitrogen | 3 | 40 80 120 |
| water | 3 | 125 250 375 |

| Number of Observations Read | 36 |
|---|---|
| Number of Observations Used | 36 |

**Dependent Variable: y**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 8 | 1.01770131 | 0.12721266 | 135.84 | <.0001 |
| Error | 27 | 0.02528594 | 0.00093652 | | |
| Corrected Total | 35 | 1.04298725 | | | |

| R-Square | Coeff Var | Root MSE | y Mean |
|---|---|---|---|
| 0.975756 | 3.499961 | 0.030603 | 0.874368 |

| Source | DF | Type II SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| nitrogen | 2 | 0.12039036 | 0.06019518 | 64.28 | <.0001 |
| water | 2 | 0.85496135 | 0.42748068 | 456.46 | <.0001 |
| nitrogen*water | 4 | 0.04234959 | 0.01058740 | 11.31 | <.0001 |

Figure 14.15: `Maestre_biomass2_new.sas - proc glm`

### Two-way ANOVA for biomass
### Data from Maestre and Reynolds (2007)
### MODEL WITH INTERACTION - USE THIS OUTPUT IF INTERACTION SIGNIFICANT

### The GLM Procedure
### Least Squares Means

| nitrogen*water Effect Sliced by water for y | | | | | |
|---|---|---|---|---|---|
| water | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| 125 | 2 | 0.009497 | 0.004749 | 5.07 | 0.0135 |
| 250 | 2 | 0.023034 | 0.011517 | 12.30 | 0.0002 |
| 375 | 2 | 0.130209 | 0.065104 | 69.52 | <.0001 |

### Two-way ANOVA for biomass
### Data from Maestre and Reynolds (2007)
### MODEL WITH INTERACTION - USE THIS OUTPUT IF INTERACTION SIGNIFICANT

### The GLM Procedure
### Least Squares Means

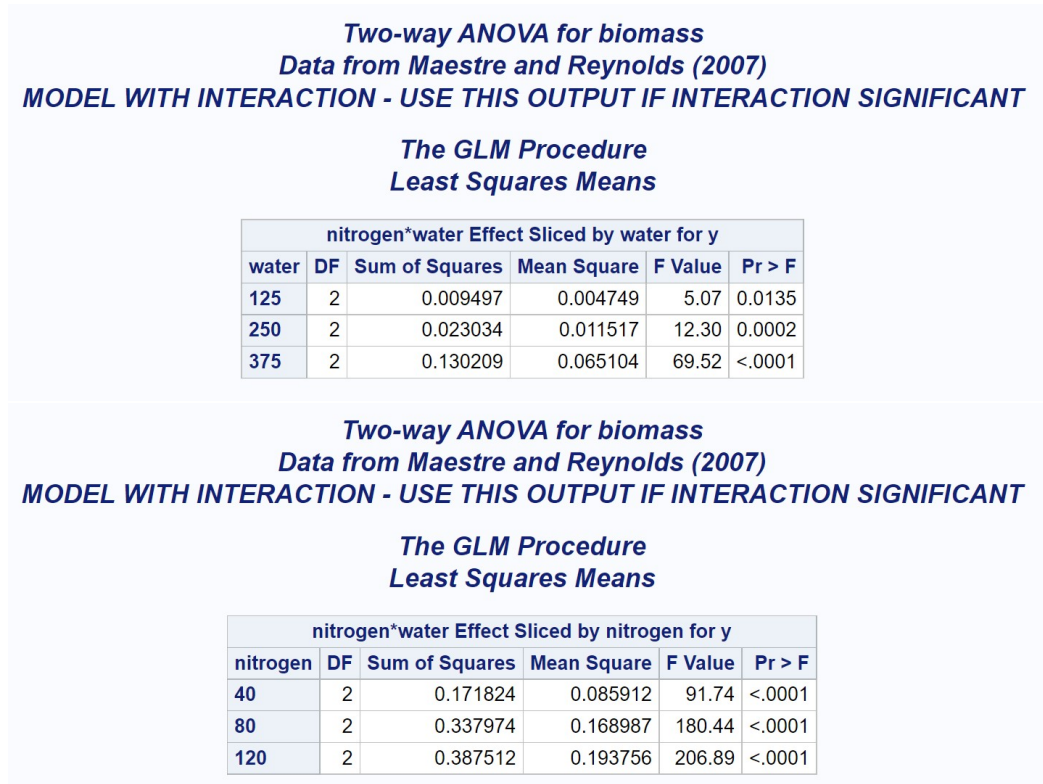| nitrogen*water Effect Sliced by nitrogen for y | | | | | |
|---|---|---|---|---|---|
| nitrogen | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| 40 | 2 | 0.171824 | 0.085912 | 91.74 | <.0001 |
| 80 | 2 | 0.337974 | 0.168987 | 180.44 | <.0001 |
| 120 | 2 | 0.387512 | 0.193756 | 206.89 | <.0001 |

Figure 14.16: `Maestre_biomass2_new.sas - proc glm`

## 14.4    Unbalanced designs and two-way ANOVA

The examples we have examined so far are balanced designs, with equal numbers of observations in each treatment combination. For these designs, the various sums of squares are independent and additive ($SS_A + SS_B + SS_{AB} + SS_{within} = SS_{total}$), the different methods of calculating the sum of squares (Type I, II, and III) yield the same results, and the resulting tests are the same. This is not the case for unbalanced two-way (or higher) designs, which occur frequently in practice. These are designs where there are fewer observations in some treatments than others, possible only a single observation. These designs can be analyzed using the same SAS procedures and programs as before, but the various sums of squares are no longer additive, and the tests are not independent (Shaw & Mitchell-Olds 1993). For this reason, if the lack of balance is severe the analysis should be interpreted with some caution.

We will use the Example 2 data set, with nine observations removed, to illustrate the analysis of unbalanced designs (see Table 14.5). The number of observations varies from $n = 1$ to 4 across treatments. These data can be analyzed using the same program as before. To show the results for both Type II and III sums of squares, the option \ `ss2 ss3` was added to the `model` statement. Examining the output (Fig. 14.17), we see that the bait × trap color interaction was nonsignificant ($F_{2,9} = 0.29, P = 0.7563$). The color effect was also nonsignificant (Type II: $F_{1,9} = 0.94, P = 0.3576$, Type III: $F_{1,9} = 0.98, P = 0.3475$), but the bait effect was highly significant (Type II: $F_{2,9} = 8.11, P < 0.0097$, Type III: $F_{2,9} = 8.15, P = 0.0096$). This is basically the same result as we obtained earlier for this study, despite the lack of balance. We can also see that the sums of squares are no longer additive. For example, with Type III sums of squares we have $SS_A + SS_B + SS_{AB} + SS_{within} = 0.9106 + 0.0549 + 0.0322 + 0.5028 = 1.5005$. This does not equal $SS_{total} = 1.4562$.

Table 14.5: Example 2 - Unbalanced design.

| Bait | Color | *T. dubius* |
|------|-------|------------:|
| FRT  | B     | 18 |
| FRT  | W     | 12 |
| FRT  | W     | 15 |
| FRT  | W     | 7 |
| FRT  | W     | 4 |
| IDT  | B     | 2 |
| IDT  | B     | 1 |
| IDT  | B     | 4 |
| IDT  | W     | 2 |
| IDT  | W     | 1 |
| IST  | B     | 2 |
| IST  | B     | 2 |
| IST  | B     | 10 |
| IST  | B     | 7 |
| IST  | W     | 4 |

**Two-way ANOVA for T. dubius counts**
**Data from Reeve et al. (2009)**

**The GLM Procedure**

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| bait | 3 | FRT IDT IST |
| color | 2 | B W |

| | |
|---|---|
| Number of Observations Read | 15 |
| Number of Observations Used | 15 |

**Two-way ANOVA for T. dubius counts**
**Data from Reeve et al. (2009)**

**The GLM Procedure**

**Dependent Variable: y**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 5 | 0.95343427 | 0.19068685 | 3.41 | 0.0526 |
| Error | 9 | 0.50280246 | 0.05586694 | | |
| Corrected Total | 14 | 1.45623672 | | | |

| R-Square | Coeff Var | Root MSE | y Mean |
|---|---|---|---|
| 0.654725 | 32.07997 | 0.236362 | 0.736790 |

| Source | DF | Type II SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| bait | 2 | 0.90584753 | 0.45292376 | 8.11 | 0.0097 |
| color | 1 | 0.05252717 | 0.05252717 | 0.94 | 0.3576 |
| bait*color | 2 | 0.03219452 | 0.01609726 | 0.29 | 0.7563 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| bait | 2 | 0.91062846 | 0.45531423 | 8.15 | 0.0096 |
| color | 1 | 0.05488645 | 0.05488645 | 0.98 | 0.3475 |
| bait*color | 2 | 0.03219452 | 0.01609726 | 0.29 | 0.7563 |

Figure 14.17: `Tdubius_bait_color_unbalanced.sas - proc glm`

# 14.5 Two-way ANOVA without replication

The designs we have examined so far assume there are multiple observations for each treatment combination, implying $n > 1$ for each group. However, it is possible to analyze studies where there is only replicate per group ($n = 1$) although this requires a change in the model. With so little data, it is not possible to estimate the interaction terms nor easily conduct a test for the interaction. However, we can fit a simplified model of the form

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}. \tag{14.42}$$

Note that the interaction term is absent. In addition, we no longer need the third subscript $k$ for the observations because there is only one observation per treatment group. One can visualize the behavior of this model using the same figures as for the two-way model with replication (see Fig. 14.3-14.5), except that the model does not incorporate interaction.

**It is important to realize that interaction could still be present in the data, even though we cannot test for it using this model.** If interaction is present it will reduce the power to detect main effects, because it adds variability to the observations in a way not accounted for by the model. Even if interaction is absent, this design will obviously have less power than a design with replication.

For these designs, we will be interested in testing whether Factor A or B have an effect on the groups means. For Factor A, this amounts to testing $H_0$ : all $\alpha_i = 0$, while for Factor B we would test $H_0$ : all $\beta_j = 0$. No test of this type is possible for the interaction.

## 14.5.1 Hypothesis testing

The sums of squares, mean squares, and other quantities for two-way ANOVA without replication are similar to those for designs with replication. We will illustrate the calculations using another data set for the insect predator *T. dubius* (Example 3, Table 14.6). This predator is most abundant during cool periods of the year in the southern USA, possibly because it cannot tolerate high temperatures. A study was conducted to see how temperature (which we call Factor A) and relative humidity (Factor B) affect the mortality rate of its eggs in the laboratory (Reeve 2000). Eggs and environmental chambers were in short supply, however, so only a single replicate was conducted at each temperature and humidity combination. Six temperatures

($15°, 20°, 25°, 30°, 35°$, and $37.5°C$) and three relative humidity treatments (55%, 75%, and 100%) were used. This corresponds to $a = 6$ and $b = 3$ in the formulas below. An arcsine-square root transformation was applied to the mortality rate observations, a common practice for data in the form of proportions.

We begin by calculating a mean corresponding to each level of Factor A by averaging across the levels of Factor B, which we denote as $\bar{Y}_{i\cdot}$. It can be calculated using the formula

$$\bar{Y}_{i\cdot} = \frac{\sum_{j=1}^{b} Y_{ij}}{b}. \tag{14.43}$$

For example, we have

$$\bar{Y}_{1\cdot} = \frac{0.379 + 0.325 + 0.615}{3} = 0.440 \tag{14.44}$$

for the first temperature treatment ($15°C$) in Example 3. The means for other temperature values are given in Table 14.6. We similarly can find means corresponding to each level of Factor B by averaging across the levels of Factor A. The general formula is

$$\bar{Y}_{\cdot j} = \frac{\sum_{i=1}^{a} Y_{ij}}{a}. \tag{14.45}$$

For the first humidity treatment in Example 3, we have

$$\bar{Y}_{\cdot 1} = \frac{0.379 + 0.439 + 0.358 + 0.466 + 0.970 + 1.571}{6} = 0.697. \tag{14.46}$$

The means for the other humidity treatments are $\bar{Y}_{\cdot 2} = 0.731$ and $\bar{Y}_{\cdot 3} = 0.719$. A grand mean $\bar{\bar{Y}}$ can then be calculated by averaging across the values of $\bar{Y}_{i\cdot}$ or equivalently by summing all the observations and dividing by their total number. It can be generally calculated using the formula

$$\bar{\bar{Y}} = \frac{\sum_{i=1}^{a} \bar{Y}_{i\cdot}}{a}. \tag{14.47}$$

For the Example 3 data set, we have

$$\bar{\bar{Y}} = \frac{0.440 + 0.502 + 0.454 + 0.521 + 0.806 + 1.571}{6} = 0.716. \tag{14.48}$$

We next develop sums of squares and means squares for this design. The difference $\bar{Y}_{i\cdot} - \bar{\bar{Y}}$ is a measure of the shift generated by Factor A in the observations, and also estimates $\alpha_i$. Squaring and summing them across all the levels of Factor A, we obtain $SS_A$. It is calculated using the general formula

$$SS_A = b \sum_{i=1}^{a} (\bar{Y}_{i\cdot} - \bar{\bar{Y}})^2. \tag{14.49}$$

$SS_A$ has $a - 1$ degrees of freedom. Its mean square is calculated using the formula

$$MS_A = \frac{SS_A}{a - 1}. \tag{14.50}$$

Note the factor $b$ in the expression for $SS_A$, which as usual scales $MS_A$ so that it estimates $\sigma^2$ under $H_0$. For the Example 3 data, we have

$$SS_A = 3 \left[ (0.440 - 0.716)^2 + (0.502 - 0.716)^2 + \cdots + (1.571 - 0.716)^2 \right] \tag{14.51}$$

$$= 3 \left[ 0.076176 + 0.045796 + 0.068644 + 0.038025 + 0.008100 + 0.731025 \right] \tag{14.52}$$

$$= 2.903298 \tag{14.53}$$

$$\tag{14.54}$$

and

$$MS_A = \frac{2.903298}{6 - 1} = 0.580660. \tag{14.55}$$

We similarly define $SS_B$ using the general formula

$$SS_B = a \sum_{j=1}^{b} (\bar{Y}_{\cdot j} - \bar{\bar{Y}})^2. \tag{14.56}$$

$SS_B$ has $b - 1$ degrees of freedom. We can then calculate a mean square for Factor B using the formula

$$MS_B = \frac{SS_B}{b - 1}. \tag{14.57}$$

For the Example 3 data, we have

$$SS_B = 6 \left[ (0.697 - 0.716)^2 + (0.731 - 0.716)^2 + (0.719 - 0.716)^2 \right] \tag{14.58}$$

$$= 6 \left[ 0.000361 + 0.000225 + 0.000009 \right] \tag{14.59}$$

$$= 0.003570. \tag{14.60}$$

and

$$MS_B = \frac{0.003570}{3 - 1} = 0.001785. \tag{14.61}$$

We now need a measure of the variability of the observations. We previously used $SS_{within}$ for this purpose, which measured the variability of the observations within each treatment group. However, in two-way designs without replication there is only a single observation in these groups ($n = 1$). If we assume there is no interaction, however, we can use an interaction-like sum of squares as a measure of variability. In particular, we have

$$SS_{within} = \sum_{i=1}^{a} \sum_{j=1}^{b} (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{\bar{Y}})^2. \tag{14.62}$$

The squared terms within this expression measure the difference between the one observation for each treatment combination and the values predicted by the model without any interaction. Note the similarity to $SS_{AB}$ for designs with replication. $SS_{within}$ has $(a - 1)(b - 1)$ degrees of freedom, and the associated mean square is defined by the formula

$$MS_{within} = \frac{SS_{within}}{(a - 1)(b - 1)}. \tag{14.63}$$

The last column of Table 14.6 shows the preliminary calculations for $SS_{within}$. Adding this column across all the treatment groups yields

$$SS_{within} = 0.131334 \tag{14.64}$$

and

$$MS_{within} = \frac{0.131334}{(6 - 1)(3 - 1)} = 0.013133. \tag{14.65}$$

The total sum of squares is given by the formula

$$SS_{total} = \sum_{i=1}^{a} \sum_{j=1}^{b} (Y_{ij} - \bar{\bar{Y}})^2 \tag{14.66}$$

and has $ab - 1$ degrees of freedom. For Example 3, we calculate that $SS_{total} = 3.038202$ with 17 degrees of freedom.

As before, we can organize the different sum of squares and mean squares into an ANOVA table. Table 14.7 shows the general layout of such a table

for two-way designs without replication. We use $F_s = MS_A/MS_{within}$ to test for the effect of Factor A. Under $H_0$ : all $\alpha_i = 0$ this statistic has an $F$ distribution with $df_1 = a - 1$ and $df_2 = (a - 1)(b - 1)$. Similarly, we use $F_s = MS_B/MS_{within}$ to test for an effect of Factor B. Under $H_0$ : all $\beta_j = 0$ it has an $F$ distribution with $df_1 = b - 1$ and $df_2 = (a - 1)(b - 1)$.

Table 14.8 shows the results for the Example 3 data set, including the $F$ statistics and $P$ values obtained using Table F. The temperature effect was highly significant ($F_{5,10} = 44.214, P < 0.001$) while humidity was nonsignificant ($F_{2,10} = 0.136, P > 0.100$). Examining the data in Table 14.6, we see that mortality rates sharply increased as temperature increased.

Table 14.6: Example 3 - Effect of temperature and relative humidity on the mortality rate of $T.$ $dubius$ eggs. Also shown are the means for each temperature level $(\bar{Y}_{i.})$ and preliminary calculations to find $SS_{within}$

| Temp. (°C) | Humidity (%) | Mortality | $Y_{ij} = \sin^{-1}(\sqrt{\text{Mortality}})$ | $i$ | $j$ | $\bar{Y}_{i.}$ | $(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{\bar{Y}})^2$ |
|---|---|---|---|---|---|---|---|
| 15 | 55 | 0.137 | 0.379 | 1 | 1 | | 0.001764 |
| 15 | 75 | 0.102 | 0.325 | 1 | 2 | 0.440 | 0.016900 |
| 15 | 100 | 0.333 | 0.615 | 1 | 3 | | 0.029584 |
| 20 | 55 | 0.181 | 0.439 | 2 | 1 | | 0.001936 |
| 20 | 75 | 0.337 | 0.619 | 2 | 2 | 0.502 | 0.010404 |
| 20 | 100 | 0.188 | 0.448 | 2 | 3 | | 0.003249 |
| 25 | 55 | 0.123 | 0.358 | 3 | 1 | | 0.005929 |
| 25 | 75 | 0.259 | 0.534 | 3 | 2 | 0.454 | 0.004225 |
| 25 | 100 | 0.205 | 0.470 | 3 | 3 | | 0.000169 |
| 30 | 55 | 0.202 | 0.466 | 4 | 1 | | 0.001296 |
| 30 | 75 | 0.321 | 0.602 | 4 | 2 | 0.521 | 0.004356 |
| 30 | 100 | 0.226 | 0.495 | 4 | 3 | | 0.000841 |
| 35 | 55 | 0.680 | 0.970 | 5 | 1 | | 0.033489 |
| 35 | 75 | 0.447 | 0.732 | 5 | 2 | 0.806 | 0.007921 |
| 35 | 100 | 0.431 | 0.716 | 5 | 3 | | 0.008649 |
| 37.5 | 55 | 1.000 | 1.571 | 6 | 1 | | 0.000361 |
| 37.5 | 75 | 1.000 | 1.571 | 6 | 2 | 1.571 | 0.000225 |
| 37.5 | 100 | 1.000 | 1.571 | 6 | 3 | | 0.000036 |

Table 14.7: General ANOVA table for two-way designs without replication, showing formulas for different mean squares and $F$ tests.

| Source | df | Sum of squares | Mean square | $F_s$ |
|---|---|---|---|---|
| Factor A | $a - 1$ | $SS_A$ | $MS_A = SS_A/(a-1)$ | $MS_A/MS_{within}$ |
| Factor B | $b - 1$ | $SS_B$ | $MS_B = SS_B/(b-1)$ | $MS_B/MS_{within}$ |
| Within | $(a-1)(b-1)$ | $SS_{within}$ | $MS_{within} = SS_{within}/(a-1)(b-1)$ | |
| Total | $ab - 1$ | $SS_{total}$ | | |

Table 14.8: ANOVA table for the Example 3 data set, including $P$ values for the tests.

| Source | df | Sum of squares | Mean square | $F_s$ | $P$ |
|---|---|---|---|---|---|
| Temperature | 5 | 2.903298 | 0.580660 | 44.214 | $< 0.001$ |
| Humidity | 2 | 0.003570 | 0.001785 | 0.136 | $> 0.100$ |
| Within | 10 | 0.131334 | 0.013133 | | |
| Total | 17 | 3.038198 | | | |

## 14.5.2   Two-way ANOVA no replication - SAS demo

We now analyze these same data using SAS. The program is similar to previous ones for two-way designs with replication, except that the interaction term needs to be deleted from the `model` statement. Because there are several levels of temperature (`temp`) and relative humidity (`rh`) in the experimental design, it seems reasonable to use multiple comparisons to compare the different groups using an `lsmeans` statement. See SAS program and output below.

Examining Fig. 14.20, we see a highly significant effect of temperature on egg mortality ($F_{5,10} = 44.31, P < 0.0001$), while the effect of humidity was nonsignificant ($F_{2,10} = 0.13, P = 0.8777$). The results are similar to the manual calculations in Table 14.6. The Tukey procedure (Fig. 14.21) finds that 37.5°C was significantly different from all the other temperatures, while 35°C was significantly different from 15°C and 25°C. No other differences were significant. There were also no significant differences among the humidity treatments. Figure 14.19 suggests that mortality was constant up to 30°C, then rapidly increased.

─────────────────── SAS Program ───────────────────

```
* Clerid_eggs_th.sas;
title "Two-way ANOVA for T. dubius egg mortality";
title2 "No replication";
data mortality;
    input temp rh mortrate;
    * Apply transformations here;
    y = arsin(sqrt(mortrate));
    datalines;
15   55  0.137
15   75  0.102
15  100  0.333
20   55  0.181
20   75  0.337
20  100  0.188
25   55  0.123
25   75  0.259
25  100  0.205
30   55  0.202
30   75  0.321
30  100  0.226
35   55  0.680
35   75  0.447
```

```
35 100  0.431
37.5 55 1.000
37.5 75 1.000
37.5 100 1.000
;
run;
* Print data set;
proc print data=mortality;
run;
* Plot means, standard errors, and observations;
proc gplot data=mortality;
    plot y*temp=rh / vaxis=axis1 haxis=axis1 legend=legend1;
    symbol1 i=j v=star height=2 width=3;
    axis1 label=(height=2) value=(height=2) width=3 major=(width=2) minor=none;
    legend1 label=(height=2) value=(height=2);
run;
* Two-way ANOVA with all fixed effects;
proc glm plots=diagnostics data=mortality;
    class temp rh;
    model y = temp rh;
    lsmeans temp rh / adjust=tukey cl lines;
run;
quit;
```

**Two-way ANOVA for T. dubius egg mortality**
**No replication**

| Obs | temp | rh | mortrate | y |
|---|---|---|---|---|
| 1 | 15.0 | 55 | 0.137 | 0.37915 |
| 2 | 15.0 | 75 | 0.102 | 0.32507 |
| 3 | 15.0 | 100 | 0.333 | 0.61513 |
| 4 | 20.0 | 55 | 0.181 | 0.43945 |
| 5 | 20.0 | 75 | 0.337 | 0.61936 |
| 6 | 20.0 | 100 | 0.188 | 0.44847 |
| 7 | 25.0 | 55 | 0.123 | 0.35833 |
| 8 | 25.0 | 75 | 0.259 | 0.53393 |
| 9 | 25.0 | 100 | 0.205 | 0.46987 |
| 10 | 30.0 | 55 | 0.202 | 0.46614 |
| 11 | 30.0 | 75 | 0.321 | 0.60234 |
| 12 | 30.0 | 100 | 0.226 | 0.49541 |
| 13 | 35.0 | 55 | 0.680 | 0.96953 |
| 14 | 35.0 | 75 | 0.447 | 0.73230 |
| 15 | 35.0 | 100 | 0.431 | 0.71618 |
| 16 | 37.5 | 55 | 1.000 | 1.57080 |
| 17 | 37.5 | 75 | 1.000 | 1.57080 |
| 18 | 37.5 | 100 | 1.000 | 1.57080 |

etc.

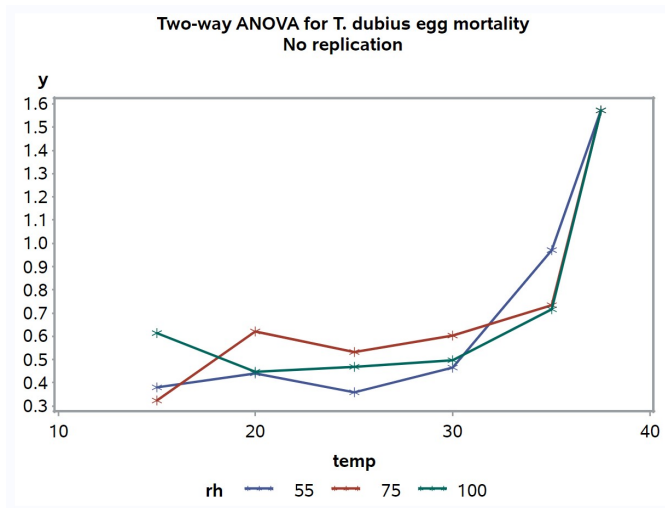Figure 14.18: `Clerid_eggs_th.sas - proc print`

Figure 14.19: `Clerid_eggs_th.sas` - `proc gplot`

**Two-way ANOVA for T. dubius egg mortality**
**No replication**

**The GLM Procedure**

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| temp | 6 | 15 20 25 30 35 37.5 |
| rh | 3 | 55 75 100 |

| Number of Observations Read | 18 |
|---|---|
| Number of Observations Used | 18 |

**Two-way ANOVA for T. dubius egg mortality**
**No replication**

**The GLM Procedure**

**Dependent Variable: y**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 7 | 2.90510830 | 0.41501547 | 31.68 | <.0001 |
| Error | 10 | 0.13098482 | 0.01309848 | | |
| Corrected Total | 17 | 3.03609312 | | | |

| R-Square | Coeff Var | Root MSE | y Mean |
|---|---|---|---|
| 0.956857 | 15.99058 | 0.114449 | 0.715725 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| temp | 5 | 2.90164653 | 0.58032931 | 44.31 | <.0001 |
| rh | 2 | 0.00346178 | 0.00173089 | 0.13 | 0.8777 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| temp | 5 | 2.90164653 | 0.58032931 | 44.31 | <.0001 |
| rh | 2 | 0.00346178 | 0.00173089 | 0.13 | 0.8777 |

Figure 14.20: `Clerid_eggs_th.sas - proc glm`

**y Tukey Grouping for LS-Means of temp (Alpha = 0.05)**

LS-means covered by the same bar are not significantly different.

| temp | Estimate |
|------|----------|
| 37.5 | 1.5708 |
| 35 | 0.8060 |
| 30 | 0.5213 |
| 20 | 0.5024 |
| 25 | 0.4540 |
| 15 | 0.4398 |

**y Tukey Grouping for LS-Means of rh (Alpha = 0.05)**

LS-means covered by the same bar are not significantly different.

| rh | Estimate |
|----|----------|
| 75 | 0.7306 |
| 100 | 0.7193 |
| 55 | 0.6972 |
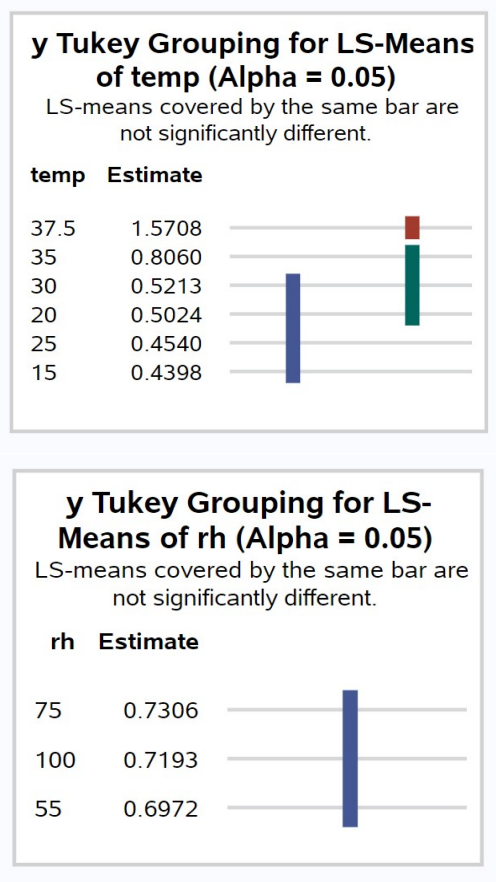
Figure 14.21: `Clerid_eggs_th.sas - proc glm`

## 14.6   Randomized block designs

Suppose that we are interested in the yield of five different strains (A, B, C, D, and E) of corn, with five replicates per strain. One possible design would be to randomly assign the strain treatments to 30 small plots scattered throughout a large field, in a completely randomized design (Fig. 14.22). The resulting data from this design could be analyzed using one-way ANOVA (Chapter 11), with strain as the treatment. One problem with this design is soil fertility, moisture, and other factors could vary across this large field. This spatial heterogeneity would make it more difficult to see any treatment effects because it would increase the variance among replicate plots.

A common two-way design, the **randomized block design**, provides a possible solution to this spatial heterogeneity problem. Suppose that soil fertility and moisture are more homogeneous on smaller spatial scales, as often seems to be true. We could then select six plots within this field, called **blocks**, and within sections of each block plant the five corn strains (see Fig. 14.23). The order of the different treatments within each block would be randomized, hence the name randomized blocks. This ensures that the sequence of treatments varies across blocks, and that each treatment has different strains for neighbors in each block. The resulting data would then be analyzed using a two-way model with a fixed treatment effect and a random block effect, which helps account and control for spatial heterogeneity in the system. The block is considered a random effect because the blocks are usually selected from a potentially large collection of possible blocks. **A statistical model with both fixed and random effects is called a mixed model.**

Another example of a randomized block design could be insect traps baited with different attractants, say A, B, C, D, and E. Different stands in the forest would be the blocks. Five traps would be deployed in each stand along a transect, with baits randomly assigned to the traps within the transect. In another type of randomized block design, the blocks are different times rather than locations in space. For example, suppose that we want to test six different diets for rearing fish in ponds, but only have six ponds available. We could randomly assign the diets to the ponds and conduct the experiment, obtaining one replicate of each treatment. We would then repeat the study several more times using the same ponds, with the treatments randomly assigned each time. Each time would be treated as a separate block in the analysis.
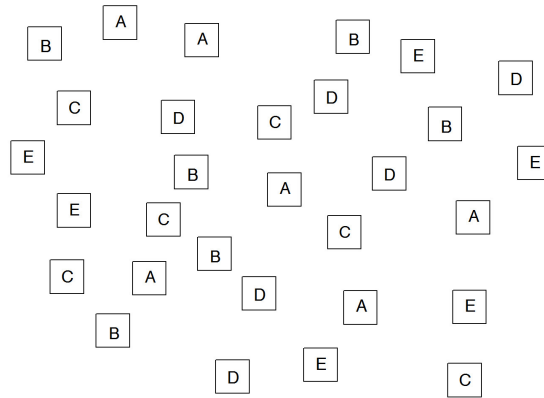
Figure 14.22: Completely randomized design with five treatments (A, B, C, D, and E) and six replicates per treatment.
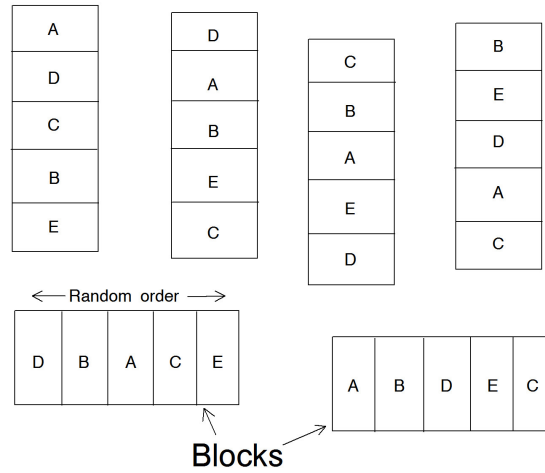


Figure 14.23: Randomized block design with five treatments (A, B, C, D, and E) and six blocks.

### 14.6.1   Randomized block models

There are two effects in a randomized block design, a fixed treatment and a random block effect, usually denoted as Factor A and B. The model commonly used to analyze these designs has the form

$$Y_{ij} = \mu + \alpha_i + B_j + \epsilon_{ij}. \tag{14.67}$$

Here $\mu$, $\alpha_i$, and $\epsilon_{ij}$ are defined as in previous models, while $B_j \sim N(0, \sigma_B^2)$. The model thus has two variance components, the variance among blocks ($\sigma_B^2$) and the variance of $\epsilon_{ij}$ ($\sigma^2$).

   Note that there is no interaction term in this model, although there could be interaction in the data. A randomized block design has just one observation per combination of treatment and block, and so there are insufficient data to estimate an A × B interaction. However, there are variants of the randomized block design that have two or more replicates of each Factor A treatment per block. In this, case, we could fit a model with interaction of the form

$$Y_{ij} = \mu + \alpha_i + B_j + (\alpha B)_{ij} + \epsilon_{ij}. \tag{14.68}$$

Here $(\alpha B)_{ij} \sim N(0, \sigma_{AB}^2)$. The interaction term in these designs is considered to be a random effect because it involves the random block effect. This model has three variance components, the interaction variance ($\sigma_{AB}^2$), the block variance ($\sigma_B^2$), and the variance of $\epsilon_{ij}$ ($\sigma^2$).

### 14.6.2   Hypothesis testing and variance components

We will use `proc mixed` in SAS to analyze the data for randomized block designs (SAS Institute Inc. 2018). The default method in SAS estimates the variance components in the model using a method called restricted maximum likelihood, or REML. This process involves separating the fixed effects from the likelihood function, then estimating the variance components of the random effects by maximizing this restricted likelihood (hence the name). Once these are determined, the fixed effects parameters are estimated and $F$ tests generated for those effects (Littell et al. 1996, McCulloch & Searle 2001). For a randomized block design, the null hypothesis tested for Factor A would be $H_0$ : all $\alpha_i = 0$. However, there is no ANOVA table nor related quantities like sum of squares and mean squares. The emphasis in `proc mixed` is on the estimation of variance components rather than tests on them, although tests can be constructed if necessary (see below).

### 14.6.3  Randomized block design - SAS demo

We will illustrate a `proc mixed` analysis for the randomized block design using a different trapping study of *T. dubius* (Reeve et al. 2009). Six different stands were located in the forest and considered to be blocks. Five traps were placed in a line at 30 m intervals within each stand, and then a bait treatment randomly assigned to each trap. There were five such treatments: blank trap (`BLANK`), $\alpha$-pinene (`AP`), frontalin + $\alpha$-pinene (`FRAP`), ipsdienol + $\alpha$-pinene (`IDAP`), and ipsenol + $\alpha$-pinene (`ISAP`). As mentioned earlier, frontalin, ipsdienol, and ipsenol are bark beetle pheromones while $\alpha$-pinene is a major component of pine resin. The number of predators caught in each trap was then counted. See SAS program with data below.

   The count data were manipulated in two ways before analysis. A log transformation was applied to predator counts to ensure the observations meet the assumptions of ANOVA (see Chapter 15). All observations for the `BLANK` treatment were also removed using the statement

```
if treat="BLANK" then delete;
```

because this treatment caught no insects. The `proc mixed` portion of the program basically implements the model for randomized block designs. We first need to tell SAS the variables categorizing the groups in the data, using a `class` statement. For the trapping study, the variables `treat` and `block` identify the treatment and block variables, so we use the statement

```
class treat block;
```

Next, recall that the randomized block model has the form

$$Y_{ij} = \mu + \alpha_i + B_j + \epsilon_{ij}. \tag{14.69}$$

Here, the SAS variable `treat` corresponds to $\alpha_i$, the fixed effect in the model, while `block` corresponds to $B_j$, the random effect. One feature of `proc mixed` is the separation of fixed and random effects in the model – all fixed effects are placed in the `model` statement while random effects are included in the `random` statement. Thus, the `model` statement for the trapping data would be

```
model y = treat / ddfm=kr;
```

while the `random` statement is

```
random block;
```

The `ddfm=kr` option specifies the Kenward-Rogers method of calculating the degrees of freedom (SAS Institute Inc. 2018), a general method for calculating the degrees of freedom that works in a variety of circumstances. An `lsmeans` statement of the form

```
lsmeans treat / pdiff=all adjust=tukey adjdfe=row;
```

is also used to compare the different bait treatments using the Tukey method. See complete program listing and output below.

From Fig. 14.27, we see there was a highly significant effect of bait treatment on the number of predators trapped ($F_{3,13.9} = 54.68, P < 0.0001$). Note the non-integer degrees of freedom for this $F$ statistic. This has occurred because the data are unbalanced (one observation is a missing value) and `proc mixed` is adjusting the test. The Tukey output (Fig. 14.28) shows a column of adjusted $P$ values, with the adjustment made according to the Tukey procedure. Adjusted $P$ values less than 0.05 are judged to be significant. We see that every pair of bait treatments was significantly different except for IDAP vs. ISAP. The graph (Fig. 14.25) and least squares means show that `FRAP` caught the most insects, IDAP and ISAP were intermediate, while AP caught the fewest.

The `proc mixed` output also provides estimates of the two variance components in the model, the block variance ($\sigma_B^2$) and the variance of $\epsilon_{ij}$ ($\sigma^2$). They are listed under the `Covariance Parameter Estimates` in the SAS output (Fig. 14.27), labeled as `block` and `Residual`, along with confidence intervals for these estimates. We see that the block variance $\sigma_B^2 = 0.3332$ was large relative to $\sigma^2 = 0.1831$. The block variance can be directly observed in Fig. 14.25 as the vertical spread between different blocks. In most cases, we are primarily interested in testing the fixed effects in the model, with the random effects and their associated variance components of less importance. They are included in the model and analysis to account and control for spatial heterogeneity in the observations. We will examine a likelihood ratio test for the block variance in the next section.

————————————————— SAS Program —————————————————

```
* TrapRCBD_clerids.sas;
title "Randomized block ANOVA for trapping experiment data";
data trapexp;
    input block $ treat $ count;
    * Apply transformations here;
    sqrtcount = sqrt(count);
    logcount = log(count+1);
    * Choose which variable is used for plots and anova;
    y = logcount;
    * Delete blank traps;
    if treat="BLANK" then delete;
    datalines;
1   AP      4
1   BLANK   0
1   FRAP    79
1   IDAP    7
1   ISAP    10
2   AP      1
2   BLANK   0
2   FRAP    124
2   IDAP    13
2   ISAP    20
3   AP      0
3   BLANK   0
3   FRAP    14
3   IDAP    .
3   ISAP    2
4   AP      0
4   BLANK   0
4   FRAP    15
4   IDAP    11
4   ISAP    7
5   AP      0
5   BLANK   0
5   FRAP    29
5   IDAP    7
5   ISAP    7
6   AP      2
6   BLANK   0
6   FRAP    70
6   IDAP    14
6   ISAP    20
;
```

```
run;
* Print data set;
proc print data=trapexp;
run;
* Plot means, standard errors, and observations;
proc gplot data=trapexp;
    plot y*treat=block / vaxis=axis1 haxis=axis1;
    symbol1 i=j v=star height=2 width=3;
    axis1 label=(height=2) value=(height=2) width=3 major=(width=2) minor=none;
run;
proc mixed cl plots=residualpanel data=trapexp;
    class treat block;
    model y = treat / ddfm=kr;
    random block;
    lsmeans treat / pdiff=all adjust=tukey adjdfe=row;
run;
quit;
```

### Randomized block ANOVA for trapping experiment data

| Obs | block | treat | count | sqrtcount | logcount | y |
|---|---|---|---|---|---|---|
| 1 | 1 | AP | 4 | 2.0000 | 1.60944 | 1.60944 |
| 2 | 1 | FRAP | 79 | 8.8882 | 4.38203 | 4.38203 |
| 3 | 1 | IDAP | 7 | 2.6458 | 2.07944 | 2.07944 |
| 4 | 1 | ISAP | 10 | 3.1623 | 2.39790 | 2.39790 |
| 5 | 2 | AP | 1 | 1.0000 | 0.69315 | 0.69315 |
| 6 | 2 | FRAP | 124 | 11.1355 | 4.82831 | 4.82831 |
| 7 | 2 | IDAP | 13 | 3.6056 | 2.63906 | 2.63906 |
| 8 | 2 | ISAP | 20 | 4.4721 | 3.04452 | 3.04452 |
| 9 | 3 | AP | 0 | 0.0000 | 0.00000 | 0.00000 |
| 10 | 3 | FRAP | 14 | 3.7417 | 2.70805 | 2.70805 |

etc.

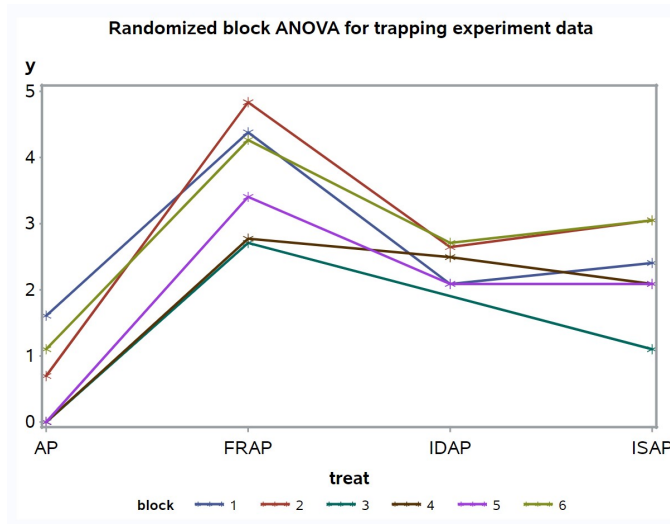Figure 14.24: `TrapRCBD_clerids.sas - proc print`

Figure 14.25: `TrapRCBD_clerids.sas` - `proc gplot`

### Randomized block ANOVA for trapping experiment data

### The Mixed Procedure

| Model Information | |
|---|---|
| Data Set | WORK.TRAPEXP |
| Dependent Variable | y |
| Covariance Structure | Variance Components |
| Estimation Method | REML |
| Residual Variance Method | Profile |
| Fixed Effects SE Method | Kenward-Roger |
| Degrees of Freedom Method | Kenward-Roger |

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| treat | 4 | AP FRAP IDAP ISAP |
| block | 6 | 1 2 3 4 5 6 |

| Dimensions | |
|---|---|
| Covariance Parameters | 2 |
| Columns in X | 5 |
| Columns in Z | 6 |
| Subjects | 1 |
| Max Obs per Subject | 23 |

Figure 14.26: `TrapRCBD_clerids.sas` - `proc mixed`

| Number of Observations | |
|---|---|
| **Number of Observations Read** | 24 |
| **Number of Observations Used** | 23 |
| **Number of Observations Not Used** | 1 |

| Iteration History | | | |
|---|---|---|---|
| **Iteration** | **Evaluations** | **-2 Res Log Like** | **Criterion** |
| 0 | 1 | 47.44629548 | |
| 1 | 2 | 38.98690259 | 0.00950955 |
| 2 | 1 | 38.96571169 | 0.00025308 |
| 3 | 1 | 38.96519017 | 0.00000021 |
| 4 | 1 | 38.96518975 | 0.00000000 |

Convergence criteria met.

| Covariance Parameter Estimates | | | | |
|---|---|---|---|---|
| **Cov Parm** | **Estimate** | **Alpha** | **Lower** | **Upper** |
| block | 0.3332 | 0.05 | 0.1159 | 3.1475 |
| Residual | 0.1831 | 0.05 | 0.09789 | 0.4576 |

| Fit Statistics | |
|---|---|
| **-2 Res Log Likelihood** | 39.0 |
| **AIC (Smaller is Better)** | 43.0 |
| **AICC (Smaller is Better)** | 43.7 |
| **BIC (Smaller is Better)** | 42.5 |

| Type 3 Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| treat | 3 | 13.9 | 54.68 | <.0001 |

Figure 14.27: `TrapRCBD_clerids.sas` - `proc mixed`

**Least Squares Means**

| Effect | treat | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|--------|-------|----------|----------------|------|---------|----------|-------|--------|--------|
| treat | AP | 0.5669 | 0.2933 | 8.59 | 1.93 | 0.0869 | 0.05 | -0.1016 | 1.2353 |
| treat | FRAP | 3.7258 | 0.2933 | 8.59 | 12.70 | <.0001 | 0.05 | 3.0574 | 4.3942 |
| treat | IDAP | 2.2417 | 0.3069 | 9.83 | 7.30 | <.0001 | 0.05 | 1.5562 | 2.9272 |
| treat | ISAP | 2.2907 | 0.2933 | 8.59 | 7.81 | <.0001 | 0.05 | 1.6223 | 2.9592 |

**Differences of Least Squares Means**

| Effect | treat | _treat | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Adjustment | Adj P | Alpha | Lower | Upper | Adj Lower | Adj Upper |
|--------|-------|--------|----------|----------------|------|---------|----------|------------|--------|-------|--------|--------|-----------|-----------|
| treat | AP | FRAP | -3.1589 | 0.2470 | 13.9 | -12.79 | <.0001 | Tukey-Kramer | <.0001 | 0.05 | -3.6892 | -2.6287 | -3.8777 | -2.4402 |
| treat | AP | IDAP | -1.6748 | 0.2630 | 14 | -6.37 | <.0001 | Tukey-Kramer | <.0001 | 0.05 | -2.2389 | -1.1108 | -2.4392 | -0.9105 |
| treat | AP | ISAP | -1.7239 | 0.2470 | 13.9 | -6.98 | <.0001 | Tukey-Kramer | <.0001 | 0.05 | -2.2541 | -1.1936 | -2.4426 | -1.0051 |
| treat | FRAP | IDAP | 1.4841 | 0.2630 | 14 | 5.64 | <.0001 | Tukey-Kramer | 0.0003 | 0.05 | 0.9200 | 2.0482 | 0.7197 | 2.2485 |
| treat | FRAP | ISAP | 1.4351 | 0.2470 | 13.9 | 5.81 | <.0001 | Tukey-Kramer | 0.0002 | 0.05 | 0.9048 | 1.9653 | 0.7163 | 2.1538 |
| treat | IDAP | ISAP | -0.04903 | 0.2630 | 14 | -0.19 | 0.8548 | Tukey-Kramer | 0.9976 | 0.05 | -0.6131 | 0.5151 | -0.8134 | 0.7154 |

Figure 14.28: `TrapRCBD_clerids.sas - proc mixed`

## 14.6.4   Likelihood ratio test for the block effect

In the preceding example, the block variance $\sigma_B^2 = 0.3332$ appeared large relative to $\sigma^2 = 0.1831$, the variance due to $\epsilon_{ij}$. The block effect was also clearly visible in Fig. 14.25. A further step would be a test of $H_0 : \sigma_B^2 = 0$ vs. $H_1 : \sigma_B^2 > 0$. If the test is significant it provides further evidence for variability among blocks in the density of insects. Littell et al. (1996) recommend a likelihood ratio test for this purpose.

We can construct this test by fitting two different models to the data, corresponding to $H_0$ vs. $H_1$. Under $H_0 : \sigma_B^2 = 0$ the statistical model for a randomized block design reduces to

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \tag{14.70}$$

because $B_j = 0$ for all $j$ under $H_0$. The statistical model under $H_1 : \sigma_B^2 > 0$ is just the full model for randomized block designs:

$$Y_{ij} = \mu + \alpha_i + B_j + \epsilon_{ij} \tag{14.71}$$

We now need to find maximum likelihood estimates of the model parameters under both $H_1$ and $H_0$, as well as $L_{H_0}$ and $L_{H_1}$, the maximum height of the likelihood function under $H_0$ and $H_1$. We would then use the likelihood ratio test statistic

$$-2\ln(\lambda) = 2\ln(L_{H_1}) - 2\ln(L_{H_0}). \tag{14.72}$$

The SAS program below finds the likelihoods for both models using `proc mixed`. Two separate calls to `proc mixed` are required, one for each model. The likelihoods are labeled `-2 Res Log Likelihood` in the output, which is almost the form required above except for the sign (see Fig. 14.29, 14.30). Examining the output, we see that $-2\ln(L_{H_0}) = 47.4$ and $-2\ln(L_{H_1}) = 39.0$. We then have

$$-2\ln(\lambda) = -39.0 - (-47.4) = -39.0 + 47.4 = 8.4 \tag{14.73}$$

How do we obtain a $P$ value for this test statistic? For any likelihood ratio test, the quantity $-2\ln(\lambda)$ has approximately a $\chi^2$ distribution under $H_0$. The degrees of freedom for the test are equal to the difference in the number of parameters for the two models ($H_1$ vs. $H_0$). There is a difference in one parameter between the two models here, because $H_1$ has the block variance $\sigma_B^2$ while under $H_0$ this is assumed to be zero. We therefore have $df = 1$, and from Table C find that $P < 0.005$. We are actually conducting a one-tailed

test, however, because $H_1$ is a one-tailed alternative. Thus, the $P$ value is half this quantity, or $P < 0.0025$. It appears the variance due to blocks was highly significant.

We can calculate the $P$ value more exactly using a simple SAS program (see below). In the data step, the program reads in the values of $-2\ln(L_{H_0})$, $-2\ln(L_{H_1})$, and *df*, then calculates the $P$ value using the SAS function probchi. We find that $P = 0.0019$ (Fig. 14.31).

———————————————— SAS Program ————————————————

```
* TrapRCBD_clerids_block_test.sas;
title "Randomized block ANOVA for trapping experiment data";
data trapexp;
    input block $ treat $ count;
    * Apply transformations here;
    sqrtcount = sqrt(count);
    logcount = log(count+1);
    * Choose which variable is used for plots and anova;
    y = logcount;
    * Delete blank traps;
    if treat="BLANK" then delete;
    datalines;
1   AP      4
1   BLANK   0
1   FRAP    79
1   IDAP    7
1   ISAP    10

etc.

6   AP      2
6   BLANK   0
6   FRAP    70
6   IDAP    14
6   ISAP    20
;
run;
title2 "H0 true - no block effect";
proc mixed cl data=trapexp;
    class treat;
    model y = treat / ddfm=kr;
run;
title2 "H1 true - there is a block effect";
proc mixed cl data=trapexp;
    class treat block;
```

```
    model y = treat / ddfm=kr;
    random block;
run;
quit;
```

**Covariance Parameter Estimates**

| Cov Parm | Estimate | Alpha | Lower | Upper |
|----------|----------|-------|-------|-------|
| Residual | 0.4925 | 0.05 | 0.2848 | 1.0506 |

**Fit Statistics**

| | |
|---|---|
| -2 Res Log Likelihood | 47.4 |
| AIC (Smaller is Better) | 49.4 |
| AICC (Smaller is Better) | 49.7 |
| BIC (Smaller is Better) | 50.4 |

**Type 3 Tests of Fixed Effects**

| Effect | Num DF | Den DF | F Value | Pr > F |
|--------|--------|--------|---------|--------|
| treat | 3 | 19 | 20.43 | <.0001 |

Figure 14.29: `TrapRCBD_clerids_block_test.sas` - proc mixed (1)

| Covariance Parameter Estimates | | | | |
|---|---|---|---|---|
| Cov Parm | Estimate | Alpha | Lower | Upper |
| block | 0.3332 | 0.05 | 0.1159 | 3.1475 |
| Residual | 0.1831 | 0.05 | 0.09789 | 0.4576 |

| Fit Statistics | |
|---|---|
| -2 Res Log Likelihood | 39.0 |
| AIC (Smaller is Better) | 43.0 |
| AICC (Smaller is Better) | 43.7 |
| BIC (Smaller is Better) | 42.5 |

| Type 3 Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| Effect | Num DF | Den DF | F Value | Pr > F |
| treat | 3 | 13.9 | 54.68 | <.0001 |

Figure 14.30: `TrapRCBD_clerids_block_test.sas - proc mixed (2)`

──────────────── SAS Program ────────────────

```
* lrtpvalue.sas;
title "P-value for likelihood ratio test";
data values;
    *Data are -2lnL values under H0 and H1, plus degrees of freedom;
    input m2lnLH1 m2lnLH0 df;
    m2lnl = -m2lnLH1 - (-m2lnLH0);
    * Find P-value;
    Pvalue = (1 - probchi(m2lnl,df))/2;
    datalines;
39.0 47.4 1
;
run;
proc print data=values;
run;
```

### P-value for likelihood ratio test

| Obs | m2lnLH1 | m2lnLH0 | df | m2lnl | Pvalue |
|-----|---------|---------|----|-------|--------|
| 1 | 39 | 47.4 | 1 | 8.4 | .001876105 |

Figure 14.31: `lrtpvalue.sas - proc print`

# 14.7    References

Cox, D. R. (1984) Interaction. *International Statistical Review* 52: 1-24.

Hurlbert, S. H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54: 187-211.

Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. (1996) *The SAS System for Mixed Models.* SAS Institute Inc., Cary, NC.

Maestre, F. T. & Reynolds, J. F. (2007) Amount or pattern? Grassland responses to the heterogeneity and availability of two key resources. *Ecology* 88: 501-511.

McCulloch, C. E. & Searle, S. R. (2001) *Generalized, Linear, and Mixed Models.* John Wiley & Sons, Inc., New York, NY.

Potvin, C. (1993) ANOVA: experiments in controlled environments. Pages 46-68 in *Design and Analysis of Ecological Experiments*, S. M. Scheiner and J. Gurevitch eds. Chapman & Hall, New York, NY.

Reeve, J. D., Rojas, M. G., & Morales-Ramos, J. A. (2003) Artificial diet and rearing methods for *Thanasimus dubius* (Coleoptera: Cleridae), a predator of bark beetles (Coleoptera: Scolytidae). *Biological Control* 27: 315-322.

Reeve, J. D., Strom, B. L., Rieske-Kinney, L. K., Ayres, B. D. & Costa, A. (2009) Geographic variation in prey preference in bark beetle predators. *Ecological Entomology* 34: 183-192.

SAS Institute Inc. (2016) *SAS/GRAPH 9.4: Reference, Fifth Edition.* SAS Institute Inc., Cary, NC.

SAS Institute Inc. (2018) *SAS/STAT 15.1 Users Guide* SAS Institute Inc., Cary, NC

Searle, S. R. (1971) *Linear Models.* John Wiley & Sons, Inc., New York, NY.

Shaw, R. G. & Mitchell-Olds, T. (1993) ANOVA for unbalanced data: an overview. *Ecology* 74: 1638-1645.

Speed, F. M., Hocking, R. R. & Hackney, O. P. (1978) Methods of analysis of linear models with unbalanced data. *Journal of the American Statistical Association* 73: 105-112.

Stewart-Oaten, A. (1995) Rules and judgments in statistics: three examples. *Ecology* 76: 2001-2009.

Winer, B. J., Brown, D. R. & Michels, K. M. (1991) *Statistical Principles in Experimental Design*, 3rd edition. McGraw-Hill, Inc., Boston, MA.

## 14.8   Problems

1. An entomologist is interested in how bark beetles respond to traps baited with two treatments, their own pheromone (P) vs. the pheromone plus a repellent chemical (PR). They also want to see if trap color (black vs. white) affects the response of the beetles. They conduct an experiment in which these two factors are randomly assigned to traps in one section of the forest, with five replicate traps for each treatment. The counts of bark beetles responding to each trap are listed below.

| Bait | Trap color | Counts for five replicate traps |
|------|------------|---------------------------------|
| P    | Black      | 138, 569, 196, 139, 726         |
| PR   | Black      | 96, 168, 25, 36, 152            |
| P    | White      | 174, 99, 293, 67, 122           |
| PR   | White      | 52, 27, 11, 57, 93              |

   (a) Write an appropriate ANOVA model for this design, and state which effects are fixed or random. Is it possible to include an interaction term in the model?

   (b) Use SAS to analyze these data using your ANOVA model, log transforming the observations. Interpret the results of all the tests. Attach your SAS program and output.

2. A research group is interested in the effects of diet and temperature on the growth rate of fish in aquaculture. They conduct an experiment with three different diet treatments (A, B and C) crossed with three rearing temperatures (15, 20 and 25°C). Two fish tanks are assigned to each treatment combination and the growth rate (g/week) determined for each tank. The following data were obtained:

| Diet | Temp | Growth rate (two tanks) |
|------|------|-------------------------|
| A    | 15   | 24.7, 22.3              |
| A    | 20   | 31.9, 28.9              |
| A    | 25   | 32.6, 31.3              |
| B    | 15   | 19.6, 14.2              |
| B    | 20   | 30.5, 26.5              |
| B    | 25   | 25.5, 32.8              |
| C    | 15   | 21.1, 21.3              |
| C    | 20   | 23.4, 23.4              |
| C    | 25   | 28.2, 25.8              |

(a) Write an appropriate ANOVA model for this design, and state which effects are fixed or random. Is it possible to include an interaction term in the model?

(b) Use SAS to analyze these data using your ANOVA model. You may use any method for dealing with interactions. Interpret the results of all the tests.

(c) Use the Tukey method to compare the different diet treatments, and then the temperature treatments. Interpret the results.