Exercise 4

PLSCI 7201

9/9/2020

Summary

In this exercise you are given two data sets. The goal is to determine if the designs are balanced and analyze the data using the approriate ANOVA approach (Type I, II, or III). Interpret the results of the experiment based on the output.

Example

In this short example I'll show you a few how to check if the design is balanced and how to fit the Type II and III ANOVA using the `car` package and the `drop1` function. The `drop1` function only does the Type III ANOVA. Remember the `aov()` function does a Type I analysis by default.

First install the car package if you don't already have it installed and load it to R. To install packages the package name must be in quotes. Once it is installed you can just call `library()` using the package name.

```
install.packages("car")

library(car)
```

I will create a toy data set and check that it is balanced. `table()` counts the number of times each combination of factor levels occur. For the data below, the first two columns of the data are our factors (A, B) and the last is our response variable.

```
### Example
set.seed(241417)
exdat <- data.frame(B = rep(c("B1", "B2", "B3"), each = 9), A = rep(c("A1", "A2",
    "A3"), 9), Y = rnorm(27, 0, 6))
table(exdat[, 1:2])
##
```

Fit a type I ANOVA using the `lm()` function.

```
summary(aov(Y ~ A * B, data = exdat))
```

Now create some unbalanced data.

```
exdat_unbal <- exdat

set.seed(42311)
exdat_unbal <- exdat_unbal[-sample(1:nrow(exdat_unbal), 3), ]

table(exdat_unbal[1:2])
```

Fitting a Type II and Type III ANOVA using `car`.

```
Anova(lm(Y ~ A * B, data = exdat_unbal), type = "II")
Anova(lm(Y ~ A * B, data = exdat_unbal), type = "III")
```

Recall from the lecture, that what sets the different ANOVA types apart is the way they are comparing models. Type III compares all possible reduced models (one fixed term dropped out) to a full model. As used below, `drop1.` does exactly this.

```
drop1(lm(Y ~ A * B, data = exdat_unbal), . ~ ., test = "F")
```

- **Q1.** What does it mean to have an "orthogonal" experimental design? Why is this important?

Dataset 1

A colleague was interested in studying the effects of nitrogen on yield. They evaluated three lines (B, C, and H) and four nitrogen levels. They asked you to analyze the data, but didn't give you much details on the experiment other than they think it followed a complete randomized design. The data is in file called "dataSet1.RDS". Use these data to answer the questions below.

- **Q2.** Is the data balanced?

Fit the appropriate ANOVA model and interpret the output. (Remember to use the right approach when interactions are significant or are approaching significance.)

- **Q3.** What type of ANOVA did you use? Why did you choose this approach? How are the models being compared with this approach?
- **Q4.** What hypothesis is being tested for the main effect of nitrogen?
- **Q5.** What can you conclude about the study?

Dataset 2

In this study we are interested in evaluating a mutant's sensitivity to different phytohormones (IAA, ABA and JA)[1]. We have three mutant lines: a wild-type (indicated by AA), a line that is heterozygous for the mutation (Aa), and a line that is homozygous for the mutation (aa). To evaluate hormone sensitivity, a batch of fifty seeds were grown in grown in petri dishes (45 dishes total) and the coleoptile length was measured for ten representative seedlings and the average was taken. The nine treatment combinations were randomly assigned to each dish. The technician could not remember if any dishes were lost due to contamination. Use the file DataSet2.RDS for this section.

[1] In a real study you would probably include a control treatment (no hormone) and possibly combinations of hormones. I'm not sure if the experiment described above would have any biological significance.

- **Q6.** Is the data balanced?

Fit the appropriate ANOVA model and interpret the output.

- **Q7.** What type of ANOVA did you use? Why did you choose this approach? How are the models being compared with this approach?
- **Q8.** What can you conclude about the study? "'