Exercises for statistical inference and stuff

Timothée Bonnet

January 15, 2019

## Contents

1 Understanding error structure

1.1 Broken experiments

* Exercise 1   A significant result?

We carry out a vaccine challenge experiment. There are three experimental groups (saline control / low vaccine dose / high vaccine dose), with 6 mice in each group. For each group all mice are housed together in the same cage (so there are three cages). All mice are challenged with *Shigella*, their symptom intensity is scored on day 8. A one-way ANOVA identifies an effect of the treatment, and pair-wise Bonferroni tests show the low-dose to be statistically significant different from the two other groups:

Mean symptom score by treatment group
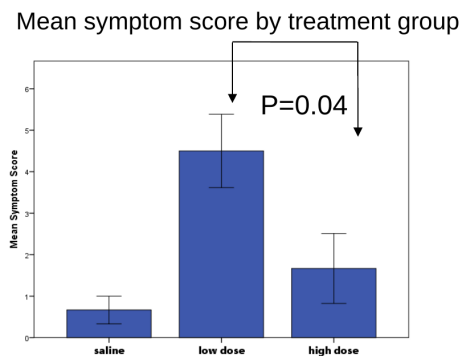
What is suspicious in the result? What part of the design may explain the result? How to improve the design (there are at least two different ways)?

* Exercise 2   A better experimental design

We study how a membrane protein intakes external molecules in frog eggs. The target molecules are radioactively labelled so we can measure intake. We created five mutant lines to test what part of the protein controls intake. We propose to measure six tubes of ten eggs every week. Each week we will test a different genotype (i.e., the control or one of the five mutants):

| Week 1 | Week 2 | Week 3 | Week 4 | Week 5 | Week 6 |

| Wild-type | Mutant 1 | Mutant 2 | Mutant 3 | Mutant 4 | Mutant 5 |

| 10 eggs/tube | 10 eggs/tube | 10 eggs/tube | 10 eggs/tube | 10 eggs/tube | 10 eggs/tube |

| x 6 | x 6 | x 6 | x 6 | x 6 | x 6 |

How much information about mutants can we extract from this experiment? How to improve the design?

1.2 Broken models?

Let's look at models with data from the an experiment similar to the mice example.

* Exercise 3   lm vs. lmer vs. lm+

Load the dataset "challenge2.csv", and test the effect of the vaccine treatment ("Group") on the immune response ("Percent.loss"). First fit a fixed-effect only linear model and/or an anova. Is there a clear effect of treatment?
Now add a random effect for Cage (e.g., using the package lme4), what happens to the effect of treatment?
Finally, fit Cage and Group as fixed effects in a simple linear model and/or anova. What happens, why?
Which of the three models should you trust or not?

1.3 Experimental design

* Exercise 4

Use a for loop and an if-else statement to do that.

2 Interpreting simple mixed models

3 Flexible variance structures