**Extra R exercises**

Once you have finished the worked examples and the exercises that are in the initial guide, choose one (or both) of the exercises.

If you don’t get to these in the course, but would like to try them out in your own time then I will be happy to answer any questions you have.

# Additional exercise 1. Power calculation

*Suppose we are designing an experiment to compare two groups of mice, with ‘weight’ as the outcome measure. The experiment will be a simple randomised study.*

*Suppose mouse weight has a standard deviation of 2g. Suppose the smallest difference between groups that we would like to be able to detect with our study is 1g.*

*How many mice do we need in order to have 80% power to detect this difference, at p<0.05?*

Look up the function power.t.test(). This is one of R’s power and sample size functions.

Use power.t.test() to find the sample size needed to detect a treatment effect difference of 1cm in two group experiment, with an outcome measure with standard deviation 2cm, using a two-sided test with p<0.05 with 80% power.

Rather than directly printing the results, assign it to an object.

Find out the class of the object you have created (using the class() function).

Find out the attributes of the object you have created (using the attributes() function).

Now print *only* the required sample size (using the $ operator).

Being able to extract just the sample size from this object is useful if, say, you want to build a table of sample sizes with varying effect size or power.

What is the smallest difference we could detect with 10 mice per group?

# Additional Exercise 2: Analysis of a randomised controlled trial

Look at the data in walkingSpeed.xlsx.

This shows the results of a randomised experiment among middle aged men and women into whether a rehabilitation treatment affected walking speed. Post-treatment times for a timed walk are given for each participant.

Use what you have learnt to load and analyse this data, in particular to test whether walking speed depends on age, sex and treatment.

Consider the following steps/questions:

Describing the data with suitable graphs and summary statistics

Mean and standard deviation of speed by sex? Or medians?

Checking whether the waking speed is normally distributed in each group

Possible transformations of the data

How well balanced were the treatment groups with respect to age?

Does treatment affect walking speed?

Is the difference statistically significant?

Are men or women faster?

Does age affect walking speed? Can you plot this relationship (using ‘plot’)