

Preliminaries:

Before you begin, ensure you have done the following (instructions are contained within the slides):

- Loaded your script from yesterday, **problems.R** into RStudio.
- Loaded your session, **problems.RData** into RStudio.
- Set the working directory as the `r_course` folder

Q1. Install and load the `pwr` package. Assuming an effect size of 0.5, use the `pwr.anova.test()` function to compute the sample size to test the difference between 4 groups, a power of 0.8 and a significance level of 0.05.

Q2. Download the **colon_cancer_p53_data.txt** file from the course website and drag to your `r_course` folder (working directory). Read in the file (`read.table()`) making sure to set the header option to `TRUE`. Examine the contents (Look at the first few rows). The data contains expression for three genes from the colon cancer data set and a label indicating if the sample was Tumour (T) or Normal (N) as well a whether the sample is a p53 wild type (W) or a p53 mutant (M). As such, there are four groups in the data, T_W (tumour, p53 wild-type), T_M (tumour, p53 mutant), N_W (normal, p53 wild-type), N_M (normal, p53 mutant).

2.1 How many groups are in the data? (Hint: `unique()` the last column)

2.1 Perform a one-way ANOVA on guanylin expression. Is there a difference between the groups? If so, perform a pairwise t-test to find which groups are different from which. Remember, to correct the p-values for multiple testing using either the FWER or FDR approach.

2.2 Perform a one-way Kruskal-Wallis test on guanylin expression? Is there a difference between the groups in this case? If so, perform a pairwise Wilcoxon test remembering to correct for multiple testing using either the FWER or FDR approach.

Statistical Programming Using the R Language

Lecture 4 – Experimental Design & ANOVA

2.3 Produce a boxplot of for the groups in TNP. Note: the boxplot function can also take a formula as an argument.

2.4 Do the results of the parametric and non-parametric tests agree?

Q3. Assuming an equal sample size per group of 9, compute the effect size (f-value) of the ANOVA performed in 2.1. Assume a significance level of 0.05 and a power of 0.8.

Q4. Once completed:

4.1 Save your script.

4.2 Save your session.

Note! Please ensure to save your work as we will use this data set continuously throughout the course.