Introduction to R

student's t-test, anova and their non-parametric equivalents

R-peer-group

QUB

February 18, 2013

Some things to begin

- Download the following .zip archive file:
  http://tinyurl.com/a3e46tp

  The archive contains all data sets for today's session, last weeks homework code, the slides for today, as well as a .Rproj file to automatically load a project for you

- When you've downloaded the archive, right click and extract to your desktop

- Navigate to the folder and double click the .Rproj file to get started

Homework review

Testing for differences

- Common questions in statistical analyses:
  - Does one set of observed data differ from another set of observed data?
  - Does a set of observed data differ from a standard distribution?
- Where data is normally distributed, generally:
  - `t-test` - Two groups
  - `ANOVA` - More than two groups
- Where data is non-normally distributed, generally:
  - `Wilcoxon test` - Two groups
  - `Kruskal-Wallis rank sum test` - More than two groups

Useful functions

```
?t.test    ?shapiro.test      ?TukeyHSD

           ?aov      ?wilcox.test

?qqnorm                ?bartlett.test

?kruskal.test                      ?var.test

?anova

                  ?qqline
```

Practical

Homework

Session-3-HW

This homework will involve carrying out the non-parametric, equivalent tests to the `t-test` and the `ANOVA` covered today

- In the folder named 'S3HW' there are two data sets
  - *Photo_1.txt*
  - *genomeSize.csv*
- You will use the *Photo_1.txt* data to carry out a non-parametric test to compare the means of paired data
- You will use the *genomeSize.csv* data to carry out a non-parametric ANOVA

You should refer back to last weeks notes for more information on the *genomeSize.csv* data. More details on the *Photo_1.txt* data are provided on the following slide

Photo_1.txt details:

- Often, before carrying out experiments on plants, it is necessary to ensure that stress effects as a result of plant translocation from the field to the lab will not impact experimental results.
- A key indicator of plant stress is low photosynthesis rate
- In the data set, photosynthesis rate was measured for a group of experimental plants, collected from the field, in two consecutive weeks after sampling
- Using this information test if there is a difference between the measurements in both weeks