Prelims Statistics and Data Analysis **TT 2019**

Lectures 1–10 = Christl Donnelly
Lectures 11–16 = Dino Sejdinovic

As usual in Prelims there will be one sheet per two lectures (eight sheets in total). So for this course this means two sheets per week – this should be about the right amount for one tutorial.

Using R or Matlab for data analysis

The main aim of this course is to teach some of the theory of statistics and data analysis. However the practical computing skills of carrying out such an analysis are important. So there are some optional questions on the problem sheets that will allow you to try basic examples of data visualization and to apply some of the techniques to real and simulated datasets. These exercises can be carried out using either R or Matlab, depending on your preference.

R `https://www.r-project.org/`

R is a free software environment for statistical computing and graphics. If you continue with statistics in Part A, then R will be recommended there (and not Matlab). If you are doing Maths & Stats or possibly interested in doing more statistics, then I would recommend you try R.

I plan to demonstrate R briefly a couple of times in lectures.

The questions on the problem sheets involving R will mostly involve you using code that you are given (either typing in brief commands, or cutting and pasting). You will not need to program in R. Getting started with R in this way should help for the future.

Note that R (and RStudio) are free, no licence is required, so anyone can install them easily.

Christl Donnelly
April 2019