

Assessed Coursework 1

This coursework is **assessed** and is worth 10% of the module mark. You should attach a yellow coursework cover sheet to your solution and submit it using the coursework collection box on A Floor of the Mathematical Sciences Building by 5pm on Monday 14th March 2016.

1. The file marij.csv contains the following variables collected in an experiment to investigate potential relationships between activity levels, body temperature and marijuana.

marij: The dose of Delta-9-THC (mg/kg) injected into study mice.

spon: Spontaneous activity as defined by the number of interruptions of a photocell beam in a clear plastic cage over a 10 minute period of time.

temp: The change in body temperature measured 60 minutes post-treatment compared to just prior to treatment.

The file Assessed_Coursework_Template.pdf contains a partially completed statistical report. Your task is to complete the statistical report following the methods stated. Your report should be a maximum of 1200 words and contain no more than 6 graphs. Please state the number of words used at the bottom of your report. Reports which exceed these limits will be penalised.

Hints and tips:

- (a) You are strongly advised to read section 3 of the lecture notes which describes the requirements of a statistical report.
- (b) You are only expected to use the statistical techniques and terminology described in sections 1-7 of the lecture notes.
- (c) The number of marks available for each section of your report are Summary [4 marks], Introduction [4 marks], Results [12 marks] and Conclusions [4 marks]. An additional 4 marks are available for the overall clarity of the writing and presentation.
- (d) The data and template report are available on the module moodle page. To read the data into R use the command `read.csv("marij.csv")`

[28 marks]

2. The Weibull distribution has cumulative distribution function (CDF) $F(x)$ given by

$$F(x) = \begin{cases} 1 - \exp(-x^2) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (1)$$

- (a) Suppose that y_1, \dots, y_n is a random sample and let $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ denote the ordered values of the y_i . Suppose that we wish to examine whether the random sample comes from a Weibull distribution using a QQ plot. Determine what value, \tilde{x}_i say, the i th order statistic $y_{(i)}$ should be plotted against. If the assumed model (1) is correct, what characteristics would you expect the plot to have? [6 marks]
- (b) Consider now the more general model

$$F(x) = \begin{cases} 1 - \exp(-\beta x^2) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2)$$

where $\beta > 0$ is an unknown parameter. Suppose that we wish to examine whether the random sample comes from model (2). If we plot the \tilde{x}_i determined using model (1) against the order statistics $y_{(i)}$, what characteristics would you expect the plot to have if model (2) is correct? How would you estimate β from the plot? [6 marks]