Department of Statistics
STATS 782 Statistical Computing
Assignment 3 (2019FC)

Total: 50 marks (including the Bonus) **Due: 3:00 pm, Wed 22 May 2019**

- Submit a hard copy with a signed, QR-coded coversheet to the SRC and a soft copy of your R code on Canvas. Other instructions might be posted on the class webpage. Ideally, the soft copy is an `.Rmd` or `.Rnw` file and the hard copy is produced by 'knitting' the file.
- Include everything in the hard copy: R code (tidied up), outputs (including error/warning messages), and your explanations (if any).
- Print some intermediate results to show how your code works step by step, if not obvious.
- Comment your code wherever appropriate, e.g., for functions, blocks of code, and key variables.
- The marker may run your R code. So include your name and ID on all files. The file names should contain your UPI. Software such as RMarkdown and knitr are recommended.
- Your mark for this assignment will depend on getting the right answer, the elegance/efficiency of your approach, and the tidiness and documentation of your code/report. The R Google Style is recommended. **Marks (up to 7) will be deducted for messy code, etc.**
- Any "**Bonus** exercises" are optional; it is possible to get full marks by completing only the compulsory parts of the assignment *perfectly*. Your mark can be improved (up to the maximum) by successfully attempting extension exercises.
- Look at the PDF version of this file as it may contain colour.
- For graphics, your R code should create a `.pdf` or `.jpg` or `.png` file that matches the original figure given for you to mimic. That is, your file should be the same dimension and colour, etc. Get your .Rmd or .Rnw file to create your graphic and then display it.

1. **[8 marks]** Consider Fig. 1, which makes one of two parallel lines of equal length look shorter. Use R to reproduce the figure as closely as possible.

Figure 1: A figure that deceptively makes two parallel lines of equal length look of unequal length.

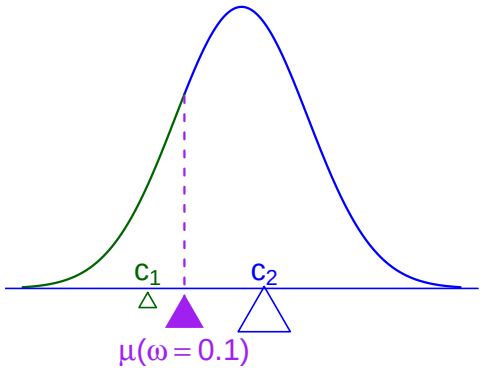

Figure 2: Illustration of the interpretation of expectiles in terms of centres of balance: the hollow triangles at positions c_1 and c_2. Here, the vertical dashed line is at the 0.1-expectile, and the solid triangle at $\mu(\omega = 0.1)$.

2. **[8 marks]** Consider Fig. 2.

   (a) Use R to reproduce the figure as closely as possible. [6 marks]

   (b) Modify the plot by choosing nicer colours, etc. in order to improve it. This should be another plot. [2 marks]

   Hints:

   - The standard normal density is the curve.
   - The positions of the triangles are $\mu(\omega = 0.1) \approx -0.862$, $c_1 \approx -1.415$, $c_2 \approx 0.342$.

3. **[10 marks]** Consider Fig. 3. Use R to reproduce the figure as closely as possible. The outer perimeter is a circle.

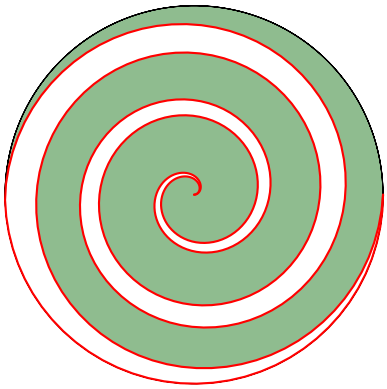

Figure 3: A figure for you to mimic.

4. **[24 marks]** Consider Fig. 4, which displays a BBC News graphic obtained from an article on recent Australian heatwaves.

   (a) The data from which the plot was made was unavailable from the BBC website. Instead, I downloaded something very similar from http://www.bom.gov.au/ and it is available to you. The data happens to come from Sydney (pronounced Seedney). Hence your plot won't be exactly the same data-wise, nevertheless you can reproduce it otherwise the same.

   Read the data into R. Then check the annual means column by computing the means over all months—you are checking the integrity of the data. Comment. [5 marks]

   (b) Try to mimic the figure in R as closely as possible, except for the BBC logo at the bottom and that the data doesn't match exactly. [10 marks]

   (c) Add in all years' data and obtain a similar plot, and use the same centring mean as in (b). Write a paragraph or two interpreting the plot, i.e., as a statistician, what do we learn from this plot? [5 marks]

   (d) Predict the annual mean temperature value in the year 2030, and try give a 95% prediction interval for this. Justify your method. [4 marks]

   (e) **Bonus**. Suppose in Fig. 4 that temperatures stabilized beyond the year 2100 and the annual mean temperature above the average is exponentially distributed with unit mean independently each year. About how many years after 2100 would somebody have to wait for so that the probability of experiencing an annual maximum greater than 6.5°C above the average is at least 5%? [3 marks]

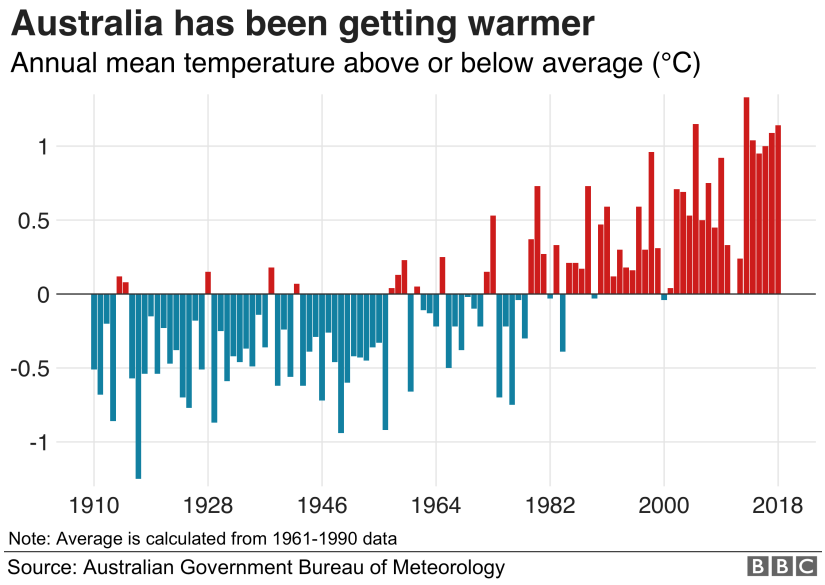

Figure 4: A figure from BBC News.