Time Series and Data Analysis

Project 2016

Instructions:

- You must submit all code written for this project in one suitable commented script.
- The examiner will run the this script to evaluate answers, so be sure to display the answers correctly.
- Discussions should be contained as comments in the file.
- Save the file using your student number, e.g. OLVCAR004.R
- Your name and student number should appear as comments on the first two lines.
- All figures must also be submitted, unless stated otherwise. Figures should be presentable, including appropriate naming of axes.

Question 1: John completely forgot to study for his exam. The exam consists of 10 multiple choice questions, each with four possible answers A, B, C, or D. John decides to choose the answers randomly.

(a) For each multiple choice question, what is the probability that John chooses the right answer? [2]
(b) Explain why the number of correct guesses is binomially distributed. [2]
(c) What is the probability that John will get exactly five answers correct? [2]
(d) What is the probability that John will get less than 5 correct guesses? [2]
(e) Generate a random sample of the number of correct guesses for 1000 such tests. Calculate the percentage of the samples with less than 5 correct guesses. Does your answer agree with (d)? [3]

Question 2: The velocities of electrons with normalized temperature *T* are normally distributed with zero mean and standard deviation $\sigma = \sqrt{T}$. Consider Plasma A with electron temperature $T_A = 1$.

(a) What is the probability that an electron in Plasma A propagates at a velocity $v_A < -1$? [2]
(b) Calculate the 70th percentile of the electron velocity distribution of Plasma A. [2]
(c) Generate a random sample of 2000 electron velocities for Plasma A. Plot a histogram of the sample. Plot the probability density function of the velocities on the same set of axes. [5]
(d) Use a quantile-quantile plot to confirm that the random sample is normally distributed. [2]

Suppose Plasma B has an electron temperature $T_B = 4$.

(e) Plot the probability density functions of the electron velocities for both Plasma A and Plasma B on the same plot. Use different colours to distinguish between the two. [4]
(f) What is the probability that an electron in Plasma B propagates at a velocity $v_B < -1$? Compare your answer to (a) to give a physical interpretation of electron temperature. [3]

Question 3: The file *Ace.Rdata* contains the plasma temperatures and solar wind velocities obtained from the ACE satellite for the year 2014.

(a) Import the data into a data frame *ACE*. Name the columns “Date”, “Time”, “T” and “V” [4]
(b) Use plots to identify missing data, and sanitise the data accordingly. You don’t have to submit these plots. [4]
(c) Create a factor *f* for the solar wind velocities with breaks at 250, 350, … [3]
(d) Calculate the mean plasma temperature for the different levels of *f*. [3]
(e) For the level 850 < V < 950, test if the plasma temperatures is significantly larger than the mean plasma temperature at a 0.01 level of significance? [3]
(f) Calculate the correlation between the plasma temperature and solar wind velocity. Test to see if the correlation is significant. [4]