Model question paper

Instructions:

(1) This paper consists of 6 questions.

(2) Three of the questions involve writing scripts in R and generating plots. Please note that for these questions, the scripts and plots should be uploaded on moodle.

(3) There are three questions for which the solution should be written using pen and paper. Note that in these cases, the final answers should be entered in moodle. If there is no entry in moodle, or, if the numbers entered in moodle conflict with the answer sheet, the answers will be discarded and will not be evaluated. For these questions, you can use R to carry out certain calculations. However, the R commands used should either be entered in moodle or should be written down in the answer sheet.

(4) The maximum time is 60 minutes.

(5) Please do not take breaks in between the examination.

(6) After the examination is submitted, within the next 30 minutes, scan and upload your written answers in the appropriate assignment. Each page of your answer sheet should have your registration number.

(7) All the best!!

Questions:

1. Consider the mother tongue data set that you are given. Write an R script that does the following:
   (a) Plot the pie chart of rural and urban speakers of two or more languages from the state of Tamilnadu.
   (b) Plot the histogram plot of two language speakers from all of India accroding to their levels of literacy.
   Note that the two plots should be one next to the other – that is, a single row with two plots. The axes should be labeled where relevant. The plots should be given their own titles. The plots should be saved in any one of the formats – pdf or jpg – and uploaded.

2. Consider the following data:

x	1	2	3	4	5	6	7	8	9
y	2.8	0.4	0.9	5.0	3.3	1.1	3.5	1.4	1.2

   Fit a linear regression line to the data; plot the data and the fitted line; plot the residuals; the plot of data and the residuals should be one below the other. Using summary command, obtain the relevant fit parameters. Comment on the significance / physical meaning of each of the terms. The plots should be saved in any one of the formats – pdf or jpg – and uploaded.

3. Write an R script to do the following: Generate 100 random variates from a log normal distribution; sample 25 data points from these variates – (a) with and (b) without replacement. Make a histograms plot of these sampled data points – one below the other. Label the plots. The plots should be saved in any one of the formats – pdf or jpg – and

uploaded.

4. A group of 5 boys and 10 girls is lined up in random order. What is the probability that the 4^{th} person is a boy? What is the probability that a specific boy is in the 12^{th} position? Solve the problem using pen and paper. Enter the final results in the text box.

5. A sample of 20 cigarettes are tested and the average nicotine content is found to be 1.2 mg. What is the 99% two sided confidence interval if the standard deviation is known to be 0.2 mg? What is the 99% upper confidence interval if the sample variance is known to be 0.04 mg? Compute a value of "c" for which you can assert with 99% confidence that the average nicotine level will be less than c. Solve the problem using pen and paper. Enter the final results in the relevant text box. If you use R for any computation, please write the relevant R command either in your answer sheet or in the text box.

6. In an experiment, 22000 healthy middle aged men were randomly split into two equal groups. Group A was given a daily dose of aspirin; Group B was given a placebo that looked and tasted like aspirin. When the experiment was stopped, in Group A, 104 men suffered heart attack and 119 men suffered stroke; in Group B, on the other hand, 189 men suffered heart attack and 84 men suffered stroke. Test the hypothesis that taking aspirin does not change the probability of heart attack. Does taking aspirin change the probability of having a stroke? Solve the problem using pen and paper. Enter the final answers in the relevant text box. If you use R for any computation, please write the relevant R command either in your answer sheet or in the text box.