INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

19th November 2019

Subject CS1B – Actuarial Statistics (Paper B)

Time allowed: 1 Hour 45 Minutes (10.15 – 12.00 Hours)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. *Mark allocations are shown in brackets.*

2. *Attempt all questions, beginning your answer to each question on a new page.*

3. *Attempt all sub-parts of the question in one document only, unless otherwise instructed to do so.*

4. *Data set file accompanying the question paper is located at: C:\Users\admin\Desktop\CS1B\filename*

5. *You need to import the same into R studio as soon as you begin the exam.*

6. *Ensure to copy and paste R codes and output at regular intervals onto the MS word file.*

7. *Please check if you have received complete Question Paper and no page is missing. If so, kindly get new set of Question Paper from the Invigilator.*

8. *Ensure that you click the "Submit" button only when you have completed the question paper and final submission has to be made.*

AT THE END OF THE EXAMINATION

Please return this question paper to the supervisor separately. You are not allowed to carry the question paper in any form with you. You are requested to save and submit the work before leaving the examination premises.

All questions need to be answered using the 'R' software, unless otherwise mentioned.

Q. 1) The probability that India will win a cricket match against South Africa is 0.7

    i) Prepare a probability distribution table of number of wins if Indians are going to play 10 matches against the South Africans. (4)

    ii) Plot a bar chart of the probabilities of number of wins from 0 to 10. (4)

    iii) Find the mean and median number of wins for India against South Africa. (4)

    [12]

Q. 2) Comment on the appropriateness of the central limit theorem by implementing the following steps:

    i) Generate a sample of 10000 random observations following Lognormal distribution with parameters $\mu = 2$ and $\sigma^2 = 0.25$ (μ and σ are the parameters of the distribution as defined in the Actuarial Tables provided) and display the first few simulated observations using the *head (...)* function. [*Use a seed value of 100 to generate random numbers*] (2)

    ii) Compute the sample mean, median and variance from the generated sample and compare the values with those of a population following a lognormal distribution with the given parameters. (8)

    iii) Treat the data generated in (i) as the population. Generate 500 different random samples of size 200 from the above population and compute the sample mean for each sample. [*Use a seed value of 100 to generate random numbers*] (5)

    iv) Plot the histogram of sample means generated from part (iii) and interpret the distribution of sample means. (3)

    [18]

Q. 3) Refer to the data file "Indices_Returns.csv" and answer the following questions:

Indices_Returns.csv file is provided in the system.

    i) Compute the pairwise Pearson correlation coefficient between the returns of 10 sectors (BM, CD, EN, FM, FI, HC, IN, IT, TE and UT) rounded to three digits after the decimal point. Display the correlation matrix in the output. (3)

    ii) Identify the pair with the highest correlation coefficient and the pair with the least correlation coefficient. (2)

    iii) Perform Principal component analysis on the returns values of the 10 sectors. (4)

    iv) How many principal components have an Eigen value of more than 1? (1)

    v) What is the approximate proportion of total variation explained by the first two principal components? (1)

vi) Compute the pair wise correlations among the 10 principal components (Round them to 3 digits after the decimal point) and display the results. What do you infer about the resulting correlations? (5)

vii) Using a scree plot comment on the number of significant components in the model. (4)

[20]

Q. 4) Refer to the data file "Indices_Returns.csv" and answer the following questions:

Indices_Returns.csv file is provided in the system.

i) Express the number of months with negative Sensex returns as a proportion of total number of months. (2)

ii) Test whether the proportion of months with negative Sensex returns is less than 50% at 95% confidence level as well as at 99% confidence level. (4)

iii) Classify the monthly returns of FI and IT sectors as follows and prepare a contingency table of counts. (5)

Monthly Return	Classification
≤ First Quartile value	Low
> Third Quartile value	High
All others	Medium

iv) Use the contingency table provided in 4(iii) above and test for the independence of monthly returns between FI and IT sector returns using an appropriate test and provide an interpretation of the test results. (5)

v) Test whether the returns of FI sector are significantly higher compared to that of IT Sector at 95% confidence level using appropriate test. (4)

[20]

Q. 5) Refer to the data file "Indices_Returns.csv" and answer the following questions:

Indices_Returns.csv file is provided in the system.

i) Load the csv file into R and create a new column called "Sensex_Direction". The value of this column will be "Positive" when the Sensex returns are positive and "Negative" when they are negative and convert the variable as a factor variable. (3)

ii) Fit an appropriate generalized linear model (GLM) to with a 'logit' link function to relate the "Sensex_Direction" with the returns of 10 sectors as a multivariate model and display the summary of the model. (6)

iii) Identify which sectors have significantly impacted the direction of Sensex returns at 95% and 99% confidence level. (4)

iv) Verify the relationship between residual deviance of the model and the Akaike Information Criteria (AIC). (3)

v) Plot the residuals of the fitted model and identify which month is the most significant outlier in the residuals. (4)

vi) Comment on the appropriateness of the model fitted. (2)

Your actuarial friend has suggested that the current model can be improved by removing the variables which do not impact the direction of Sensex returns at 95% confidence level and refitting the GLM with '*logit*' link function.

vii) Update the model fitted in (ii) above, as suggested by your friend and display the summary of the model. (4)

viii) Compare the models in (ii) and (vii) using an appropriate test and comment on the difference in the residual deviances between the two models. (4)

[30]

****