



INSTITUTE OF ACTUARIAL
& QUANTITATIVE STUDIES

Subject:

**SRM1 + R
Project**

Chapter:

Category:

Project – Statistical & Risk Modelling – 1 and Modelling in R

Project Topic: Graduation

Deadline – 15th May 2024.

Description: The following project needs to be completed in R, with your answers to the subjective questions written in a Word file. Submit a zip file (that contains both the above submissions).

The file you submit should be named as follows:

MSC SEM2 SRM1 + R PROJECT ROLL NO ____.

The project refers to Graduation and the statistical tests that can be performed for graduation. You are given enough hints to understand how various computations can be achieved. You may, however, use your own formulas/codes to achieve the results, provided they form a logical flow.

The question and assessment criteria are given below. The project is worth a total of 100 marks.

Question:

Task 1

You have been provided with a data file Graduation.csv, which includes the following columns:

AGE - Age x

ETR - Central exposed to risk for age x last birthday.

DEATHS - Number of deaths recorded at age x last birthday.

CRUDE - Crude mortality rate for age x last birthday.

GRADUATED - Graduated mortality rate for age x last birthday.

EXPECTED - Expected number of deaths at age x last birthday based on the graduated rates Individual.

ZX - standardized deviation for age x.

1) Read in the data file as a data table & fill in the entries for the CRUDE column in your table. (5)

2) Use Gompertz law to fill entries in the GRADUATED column in the data. (10)

(Hint: You may look up the function 'lm' for fitting the gompertz law to the data and the function 'coef' and 'as.numeric' to estimate the parameters B and C)

PROJECT

3) Check for smoothness by applying the third differences to the crude and graduated rates and comment on your results. (15)

(Hint: You may use the following function to calculate the 1st order differences for a vector-

1diff = function (x) x[-1] - x[-length(x)])

4) Calculate the values for EXPECTED and ZX columns in the table.

Hence, perform a chi-squared test to check goodness of fit between DEATHS and EXPECTED. You should specify the degrees of freedom used. (12)

(Hint: You may use the in-built chisq.test() function for this or directly calculate the chi-square value using ZX)

5) a. Perform the standardized deviations test on the individual deviations, and comment on the following: (12)

- i. Overall shape
- ii. Absolute deviations
- iii. Outliers
- iv. Symmetry
- v. Final conclusion about Null hypothesis

5) b. Perform the Signs test and give your conclusion (12)

(Hint: Since m is large, you may use the normal approximation along with continuity correction for this test. You may look up the function qnorm and pnorm in R for this)

5) c. Perform the Cumulative deviations test for the entire age range and give your conclusion. (12)

5) d. Perform the Serial correlations test and give your conclusion. (12)

(Hint: You may use the function 'cor' to compute the correlation between the 2 variables. You may specify the variables using following functions:

$Z1 = ZX[1 : \text{length}(ZX)-1]$ and $z2 = ZX[2 : \text{length}(ZX)]$

Task 2

Given the data 'Claims.RData', find the statistical distribution using MLE as method of estimation.

Students are expected to conclude basis appropriate charts / hypothesis test. (10 marks)

Assessment:

You have to submit the R script along with the word file for your answers.

In R, you should follow good coding practices by making suitable no. of variables for computations, not repeating/overwriting, having a logical flow of the code, and putting comments wherever required.

In word, you are expected to prepare an answer sheet to answer all the questions mentioned above and giving extracts/values from the R output wherever required.

You will be marked based on the above mark allocation where your code, result, accuracy and logic of the comments and conclusions will be tested.

Marks will be awarded based on the overall presentation of your R script and your word file, along with punctuality in the submissions.