

## MTH6157 Survival Models - R Coursework Question from 2022

Before you start this question, you will need to download the .csv file “Week9\_dataset.csv” from QM Plus and save it in your workspace.

An insurance company currently uses the force of mortality from the AM92 standard table for term insurance premium calculations. The actuary wishes to compare the mortality experience in 2021 with the rates used in premium calculations. `Week9_dataset.csv` contains 7 columns of data:

The first column contains ages 60 to 69. The next five columns are the number of term insurance policies for policyholders age  $x$  nearest birthday ( $x = 60, 61, \dots, 69$ ) at three-month intervals from 31/12/2020 to 31/12/2021. The seventh column has the number of policyholder deaths at age  $x$  nearest birthday in 2021. The final column is the AM92 table force of mortality by age.

- (a) Import the data from `Week9_dataset.csv`, save it in a dataframe called `insurance` and display the first 6 rows of data in R. [2]
- (b) Construct R code to calculate the central exposed-to-risk for policyholders age  $x = 60, 61, \dots, 69$  and store it in a vector called `EtOR`. [5]
- (c) What assumptions have you made in your calculation in (b) above? [3]
- (d) Calculate the observed forces of mortality in 2021 at ages  $x = 60, 61, \dots, 69$  assuming a two-state model for mortality and store these in a vector called `mu`. [3]
- (e) Explain how your calculation in (d) above would be different if deaths only had been recorded as age last birthday instead of age nearest birthday. *You do not need to complete this calculation in R, just write about how it would be different.* [4]
- (f) Plot the observed and AM92 table forces of mortality on a graph with legend, title, axis labels and two different colours for the two forces of mortality. [4]
- (g) What conclusions can you draw from the graph in (e) above? [2]
- (h) Test the overall goodness of fit of 2021 observed mortality to the standard table used in premium calculations using R code. In your MS Word document include the hypothesis being tested, the degrees of freedom and the results of the test as well as the R code and R output. [9]
- (i) What conclusions should the actuary make from the graph in (e) and the statistical test in (h) about the suitability of the AM92 standard table for term insurance premium calculations? Explain what further analysis or information she might need. [8]