Description

Assignment in Stochastic Modelling   This task requires you to fit a Markov chain model to simulated insurance claims data. The data are in the file 'Classification Scheme Data.csv' (Attached). The Mastodon Insurance Company studies a cohort of 600 drivers, who were all below 25 years old at the start of the study. In each year the number of claims made by every driver was noted. Mastodon operates a classification scheme with six discount levels from level 0 (no discount, ie the driver pays full premium) to level 5 (50% discount), with a 10% increase in discount at each step. A policyholder who makes no claims in a year moves up one level (unless already at level 5); a policyholder who makes 1 or more claims moves down one level (unless already at level 0). Before the study began, the drivers were categorised using variables such as age, gender, zip code, and miles driven per year. The categories reflect Mastodon’s expectation of the level of risk associated with that driver: •    Category A — very low risk, ie the best drivers •    Category B — low risk •    Category C — medium risk •    Category D — high risk, ie the worst drivers The data set The data set supplied is a .csv file Attached. The top row of the data set provided consists of headers. The category in which the driver was classified is the first column of the data set, which has the header "Category". The discount level in which the driver was located at the start of the study is the second column, which has the header "Initial". The remaining 15 columns give the number of claims in each of the years 1 through 15 of the study. Notation Denote by Xi,n the discount level in which the ith policyholder finds themselves at time 0, 1, 2, ..., 16 (time measured in years). I will refer to this as matrix X below. The task The aim of this task is to investigate 3 themes and to write up your findings in the form of an Executive Summary. You should aim for a length of 500 words, suitably illustrated with one or two graphs produced using a statistical package (R is fine) or spreadsheet. When you submit you should upload two files to Blackboard: 1.    The report, as a Word document 2.    The code and/or spreadsheet you used to generate your results, which must be appropriately documented so that a reader can follow what you did. If using R, this should be in the form of a file with a .r extension, not the output from an R-Studio session. The 3 Themes   Theme 1 Assume that the number of claims made by a customer each year forms a sequence of iid random variables, with common distribution being Poisson. You will investigate three possible cases regarding the Poisson parameter (λ).

Order 1560978