THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

MA424 Modelling in Operations Research
Katerina Papadaki

Simulation: Project 2021 - 2022

Carsafe is a small car insurance company with a fairly simple insurance model. They charge all customers the same fixed amount M per month independent of the age or type of the car. At the moment they have n_0 customers and their analysts, after looking at past data, have deduced that customers join Carsafe according to a Poisson process with rate λ per month. The analysts have also figured out that each customer leaves Carsafe according to a nonhomogeneous Poisson process with rate $\mu(t)$ per month, where t is the time that elapsed from the current time. Each Carsafe customer has an accident according to a Poisson process with yearly rate α; from those customers that have an accident there is a probability 0.6 that they will claim and the claim amount is $300X + 500$, where X is a discrete uniform random variable over the set $\{0, 1, \ldots, 10\}$. Note that after a customer has an accident, they continue having accidents according to the Poisson process until they leave Carsafe.

1 Will Carsafe survive the next 12 months?

The company has initial capital c_0 and they are worried that capital will get lower than a threshold C, in which case they are in danger of going bankrupt. They want to estimate P, the probability that in the next 12 months Carsafe's capital will go below C.

(i) Build a simulation model to help Carsafe estimate P. When you build your simulation model define your variables, events, event lists, output variables. Write down the pseudocode of each event case as we did in lecture. Use $K = 1000$ iterations and the following data: $n_0 = 0$, $c_0 = £50,000$, $M = 300$, $C = £30,000$, $\lambda = 3$, $\mu(t) = \frac{6}{2+t}$, $\alpha = 1.2$. Comment on your results.

(ii) Let C_{12} be the capital after 12 months. Find an estimate of $E[C_{12}]$ for which you are 90% confident that it is within £500 of its true value. In how many iterations did you get this value and what is the standard deviation of the estimator?

(iii) Give an interval centered around the estimate given in (ii) above where we are 95% confident that the true value of $E[C_{12}]$ is within this interval.

2 Two other models to simulate claims.

The analysts at Carsafe realised that the model described above for estimating the claim amount (given that a customer will claim) is too simplistic. Because there was some disagreement between them about what the claim amount should be like, they split into two teams and studied the problem

separately. Team 1 deduced that the claim amount was (as before) the discrete random variable $300X + 500$, but where X takes values in $\{0, 1, \ldots, 10\}$ according to $P(X = i) = Se^{-i}$, where S is a normalising constant. Team 2 deduced that the claim amount was a continuous random variable $10200Y$, where Y takes values in the interval $(0, 1)$ and has density $f(y) = Wye^{-y^2}$, where W is a normalising constant.

(a) For team 1 you are advised to use the acceptance rejection method to generate the claim amount.

(i) Show analytically that you do not need to calculate the value of S to use the acceptance rejection method. Explain clearly what proposal distribution you would use and why this distribution is appropriate.

(ii) Without running an R script calculate analytically the expected number of iterations that you will need to run the acceptance rejection method before it accepts a value. Explain how you found this.

(iii) Write an R script to simulate the distribution of the claim amount proposed by team 1 using the acceptance rejection method.

(iv) Use 10000 simulated values to check graphically that the distribution is simulated correctly. Explain how you did this. Also verify that the average number of iterations that you run the rejection method is close to your answer in (ii).

(b) For team 2 you are advised to apply Markov Chain Monte Carlo (MCMC) to simulate the claim amount.

(i) What MCMC algorithm would you use and why?

(ii) What proposal chain would you use? State why this proposal chain is appropriate.

(iii) Write an R script to simulate the distribution of the claim amount proposed by team 2 using MCMC.

(iv) Simulate 10000 values of this distribution to check graphically that the distribution is simulated correctly. Explain how you did this.

3 Reduce the variance of the estimator.

The CEO of Carsafe is very happy with your estimate of $E[C_{12}]$ but is worried about the variance of the estimator. She is wondering whether there is another estimator of smaller variance. Write a paragraph proposing a variance reduction estimator (no more than one) and explain why it would reduce the variance. (Do not perform any simulations).

Deliverables and report content - final project part 2

You are expected to write a report helping Carsafe with all their questions above. The report should be accompanied with the corresponding R scripts. The report should be divided into parts where each part corresponds to a question posed above. In your answer for each part of a question you should include both technical parts and explanations of these technical parts as well as offer intuition where appropriate.

Some further rules and pointers:

1. For generating random variables, you are only allowed to use the function **runif(n)** for n a positive integer. If you use any other function to generate random variables you will be penalised.
2. Start every R script with the command **set.seed(1)**.
3. You may use R code from some of the R scripts that I provided for the lectures or the exercises. If you do that, you need to write this clearly in your report otherwise this will be considered **plagiarism.**
4. Each R script should have detailed comments and explanations so that someone who does not understand R can figure out what the code does at each step.
5. The technical part of the report should explain what each corresponding R script does so that someone that has not seen the R code can read the report and understand what you have done.
6. Use a separate R script for each question part. If in one question you are using the same R script from a previous question but with some modifications, then make another copy of the R script and name it with the corresponding question number.
7. **Do not use an executive summary or appendices.**
8. Include all necessary graphs in the report. Make sure the graphs is at the correct position in the report.
9. The report should be typed and in pdf format.
10. Do not copy the R code in the report. Submit the R files along with the report.

Guidelines - final project part 2

(a) The deadline for submitting this project is Tuesday 12th of April 2022 at noon UK time. The pdf file of the report, and the R scripts should be contained in one .zip file or .rar file, to be uploaded on the moodle MA424 page anonymously - **only with your exam candidate number** - by the deadline. Name your zip. file with your exam candidate number only. There will be separate submission links for the Mathematical Programming and Simulation parts of the project. Please take care to submit the correct part under the correct link.

(b) On the Simulation part of the project, you are expected to work on your own. **You are not allowed to discuss your project with other class mates.** All components of the report must be written by each student individually. You may not seek advice from anyone else other than clarification from the MA424 lecturers.

(c) The R code should also be written by each student individually.

(d) The report should not exceed 8 pages. It should be with font 11pts, and single spacing. Note that in part 2 of the project you are not allowed to have appendices.

(e) Any part of the project submitted without all associated R files will be given a **mark of zero**. If an answer is given in the report that cannot be confirmed by your code, then this will be considered as **plagiarism**.

(f) You are allowed to ask only clarification questions to the MA424 lecturers.

(g) The clarification questions should be asked on the Moodle forum so that the answers are visible to all MA424 students.

Marking scheme - final project part 2

R code and R code quality of comments	40
Quality of explanation of solution in report	40
Organization/Presentation	20
Total	**100**

- R code and R code quality of comments. The correctness of the code, the use of appropriate R functions as described above, the quality of the comments that make it easy for someone who does not know R to understand what your code does.

- Quality of explanation of solution in report. The correctness of the solution and analysis, the quality of the explanation of the R scripts, and the quality of the explanations of the solution.

- Organisation/Presentation. A report should be clear, concise, well organised, well formatted and well presented with appropriate use of figures and tables, and well commented. It should also follow the instructions in terms of type of files, length, font size, and spacing.