MA424 Modelling in Operations Research
Katerina Papadaki

Simulation Project 2022 - 2023

Due to rising energy prices, the government of World Kingdom (WK) encourages its citizens to install solar panels on their roofs. They offer full or partial funding of solar panel costs to support the initiative. They created a government department, Solar Panel Funding (SPF), that processes these applications. However, complaints were made by citizens about the times spent inside the SPF offices when applicants were dropping their applications for processing. You are hired to investigate this.

The SPF offices are open on a weekday 9-5pm. Applicants arrive during these hours at the offices according to a Poisson process with rate λ per hour. As the application forms for funding are quite complex and there is a high chance the applicants made a mistake, the government recommends an application checking services (ACS) that the applicants can utilise for a small fee. SPF estimates that there is a probability p that an applicant used ACS.

When an applicant that has used ACS arrives at the SPF offices, they go straight to a server S_F that processes their application. They queue to see S_F in a FIFO fashion and when service completes they leave the offices with their application successfully submitted. If an applicant has not used ACS, they queue (FIFO) to see server S_C that will check their application and correct any errors. There are two outcomes from the S_C server: either they are satisfied with the application (with probability q) and they send the applicant to queue at S_F, or they are not satisfied and they send them to another server S_W to queue (FIFO) for an interview. When server S_W completes the interview, either they are satisfied (with probability r) and they send them to S_F, or they think they don't satisfy the criteria and they reject them, in which case they leave the offices.

The doors to the offices close to incoming applicants at 5pm every day, however, the applicants that are already inside the offices will be served as normal. When all applicants leave, then the offices close. The servers S_F, S_C, S_W have service times that follow exponential distributions with rates μ_F, μ_C, μ_W per hour respectively.

1 Find applicants' expected waiting times.

The head of the department wants to know the waiting times of applicants. Let T_a be the random variable that measures the time spent by an applicant in the SPF offices. And let T_c and T_r be the random variables for time spent by completed customers (that successfully submitted their application) and the time spent by rejected customers.

(i) Build a simulation model to help SPF estimate $E[T_a]$, $E[T_c]$, and $E[T_r]$. 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 $K = 500$ iterations and the following data: $\lambda = 8$, $\mu_F = 6$, $\mu_C = 5$, $\mu_W = 4$, $p = 0.5$, $q = 0.6$, $r = 0.4$. Report the estimates that you found for $E[T_a]$, $E[T_c]$, and $E[T_r]$. (*Hint: your state variable should contain information of the position of all incoming applicants that have entered the system so far, i.e. since the opening of the office at 9am.*)

(ii) Find an estimate of $E[T_a]$ for which you are 95% confident that it is within 10 minutes 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 90% confident that the true value of $E[T_a]$ is within this interval.

(iv) Use $K = 500$ iterations to simulate the percentage of customers completed and percentage of customers rejected. Do the same analytically and compare your answers. How many iterations do you need to run the simulation to get very close to the analytical answer?

2 Improve the estimator of $E[T_a]$.

How would you use variance reduction to improve the estimator $E[T_a]$. 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).

3 Distributions of servers S_C and S_W.

SPF informs you that the distributions for servers S_C and S_W are not as given in part 1. For S_C, the density function is $f(x) = \frac{1}{2}(1 + x)e^{-x}$, $0 < x < \infty$, where x is in minutes. For S_W, the distribution function (CDF) is $F(x) = 1 - e^{-3x^2}$, $0 < x < \infty$, where x is in hours.

(i) For each server, use a different method to simulate the service times. Justify why you used that method. Write two R functions, one for each server, that simulate the service times.

(ii) Generate 10000 iid values from each R function and plot their histogram with the densities of the distributions to show that they have indeed been simulated correctly (multiply the densities with a constant so that they align with the histogram and explain how you got that constant).

Deliverables and report content

You are expected to write a report helping SPF 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

(a) The deadline for submitting this project is April 17th 2023 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) For the simulation 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.

(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

$$Q1: \ 50 \ points$$
$$Q2: \ 10 \ points$$
$$Q3: \ 40 \ points$$