Simulation Project

Background

A company has an IT help desk for their employees with only one support person answering the phone. As many employees complain about getting a busy tone too often and spending too much time trying repeatedly, the company plans to introduce a delay line. The number of places in the delay line should be limited such that the average service time of a full queue does not exceed 30 min. Both arrival and service times can be considered being exponentially distributed (M). The steady-state performance of the existing system without delay line and of the proposed system (with delay line) are to be studied by means of quantitative simulation, which requires a sound confidence analysis. Based on those simulations, decisions regarding the dimensioning will be based.

Tasks

For sake of simplicity, we always (for all tasks) throw the first batch to take some warm-up into account.

1. Investigation of the existing system, and comparison of confidence intervals.

Use a system with the following properties:

- no delay line (queue, $K = 0$)
- use $n = 100$ batches, and
- $k \in \{250, 1000, 4000\}$ arrivals (do not bother with correlation in this task)

and estimate the *probabilities* that a caller gets a busy tone (= lost call) in the following cases:

a. Mean inter-arrival time = 5 min, mean service time = 5 min
b. Mean inter-arrival time = 5 min, mean service time = 4 min
c. Mean inter-arrival time = 5 min, mean service time = 3 min

In particular, perform the following tasks for a, b and c:

• Sketch the 95 %-confidence intervals (lower bound; middle; upper bound) as reported by the simulation program and compare them with each other. To which extent do they overlap? How about their size, especially with regard to *k*?

• Calculate 90 %- and 99 %-confidence intervals from the 95 %-confidence intervals (HINT[1]). Sketch and compare them with each other and with the corresponding 95 %-confidence intervals.

d. For each case a, b and c (to be called 1.d-a, 1.d-b and 1.d-c), run $n = 400$ batches, with $k = 1000$ arrivals and sketch 95 %-confidence intervals as before. Compare this with the 95 %-confidence intervals sketch of the case $n = 100$ batches and with $k = 4000$ arrivals.
Is there any significant difference to these sketches (e.g. 1.d-a and 1.a)? Why or why not?

Explain and motivate!

[1] HINT: There is a very simple way to do this if you just look at the z1-a/2-values for the different confidence levels ☺ see lecture and homework.

2. Condition handling.

Calculate the queue size K (NOTE: has to be an integer!) in the following cases such that the average service time of a full queue does not exceed 30 min:

a. Mean inter-arrival time = 5 min, mean service time = 5 min ➔ K_a = …
b. Mean inter-arrival time = 5 min, mean service time = 4 min ➔ K_b = …
c. Mean inter-arrival time = 5 min, mean service time = 3 min ➔ K_c = …

As the results will be needed later on, please confirm them with your supervisor.

3. Autocorrelations and batch sizes.

Estimate the lag-1 autocorrelations between the simulation results by performing n = 1000 batches, each with $k \in \{10, 100, 1000, 10000, 100000\}$ arrivals, for the following cases:

a. Mean inter-arrival time = 5 min, mean service time = 5 min, queue size = K_a
b. Mean inter-arrival time = 5 min, mean service time = 4 min, queue size = K_b
c. Mean inter-arrival time = 5 min, mean service time = 3 min, queue size = K_c

Which tendencies can you observe? Which batch size (as small as possible) provides the best lag-1 autocorrelation? Explain and motivate why this is the best one, according to theory and according to your opinion.

4. Parameter study and relative half-sizes of confidence intervals.

Estimate the call loss probability as a function of the queue size K for the following cases using n = 100 batches with k = 100000 arrivals each.

a. Mean inter-arrival time = 5 min, mean service time = 5 min,
queue size $\in \{0, 1, \ldots, K_a\}$
b. Mean inter-arrival time = 5 min, mean service time = 4 min,
queue size $\in \{0, 1, \ldots, K_b\}$
c. Mean inter-arrival time = 5 min, mean service time = 3 min,
queue size $\in \{0, 1, \ldots, K_c\}$

In each of those cases, plot

- The estimated call loss probability as a function of the queue size K
- The relative half-size of the 95 %-confidence interval as a function of the queue size K

Which tendencies can you observe?

Explain and motivate!

5. Quality-assured simulation.

For each of the following cases, determine the number of batches ($n \geq 40$)[2] such that the relative half-size of the 95 %-confidence interval amounts to 1 %:

a. Mean inter-arrival time = 5 min, mean service time = 5 min, queue size $\in \{0, 1, \ldots, K_a\}$

b. Mean inter-arrival time = 5 min, mean service time = 4 min, queue size $\in \{0, 1, \ldots, K_b\}$

c. Mean inter-arrival time = 5 min, mean service time = 3 min, queue size $\in \{0, 1, \ldots, K_c\}$

How does the number of required batches n change with respect to the queue size in each of these cases? Explain and motivate!

6. Counter-checking.

For the simulations with the maximal queue size

a. Mean inter-arrival time = 5 min, mean service time = 5 min, queue size = K_a

b. Mean inter-arrival time = 5 min, mean service time = 4 min, queue size = K_b

c. Mean inter-arrival time = 5 min, mean service time = 3 min, queue size = K_c

cf. Task 5, calculate the sum of the reported loss ratio and the reported server load and compare it to the offered load (= mean service time / mean arrival time). What do you observe? Explain and motivate!

7. Different distributions.

Replace the input and output distributions and consider the following systems with mean inter-arrival time = 5 min, mean service time = 4 min, n = 100 batches and k = 10000 arrivals:

- $D/D/1/K_b$
- $D/M/1/K_b$
- $D/U/1/K_b$
- $M/D/1/K_b$
- $M/M/1/K_b$
- $M/U/1/K_b$
- $U/D/1/K_b$
- $U/M/1/K_b$
- $U/U/1/K_b$

Sort the systems according to loss performance (do not forget to specify confidence intervals!) and try to explain why you see that particular ordering. Which one is best, which one is worst and why (NOTE[3])? Explain and motivate!

---

[2] For sake of simplicity, we use the factor 1.96 even for "sufficiently large n" (about 40).
The following approximation can be used as well for $40 \leq n \leq 100$: $\boldsymbol{z \approx 2.06 - 0.001\, n}$

[3] NOTE: That a system is best because it gets the best result is only an observation of a fact. Try to "look behind" the result, and explain why you get the best result in that particular case.

8. Dimensioning.

Obviously, the mean service time has an impact on the call loss probability. Viewing your results from Task 4, your boss would like to know by how much the service time should be reduced from the current average of 5 min in order to yield a call loss probability of 5 %. Regarding the latter, he expects a confidence interval of ±0.5 % of the mean value. Given you could use k = 100000 arrivals, find out

a. which mean service time (as close to 5 min as possible) provides you with an estimated loss of 5 %; observe that the average queuing time given a full buffer (= K x mean service time) shall not exceed 30 min; (HINT[4])

b. how many batches you need to use in order to reach a relative half-size of the confidence interval of 0.5 %;

c. the estimation of the lag-1 autocorrelation between the results.

d. Regarding the latter (c), are your confidence intervals too large or too small as compared to what they should be in the uncorrelated case?

[4] HINT: There are several parameters you can change. Try them all. And remember that you should get as close as possible to a service time of 5 min.

Hints

- You can run the program
  - In the Visual C++ environment (which gives you the possibility to change the code according to your needs).
  - In a Command Line Interface (CLI) – change to the directory in which the program mm1.exe is found, and type mm1.

  Important: Don't just click on the mm1.exe file, as the result won't be visible in the end.
- With the > operator, you can write the output of the simulation program directly to a file. It is however mandatory to start the program in a CLI.
- In the same way, you can read your input data from an input file using the < operator. The input file contains only the data that you have to insert via the keyboard, nothing else (because then, the software may hang and fill your account). The program repeats the input so that it is visible in the output as well. An example for an input file is as follows:

  123457
  M 1.25
  10
  M 1.00
  1 40
  1000

  - The first row contains the seed (0 is treated at the standard seed 123457).
  - The second row contains the source parameters (M process, interarrival time).
  - The third row contains the queue size.
  - The fourth row contains the server parameters (M process, service time).
  - The fifth row contains whether a warm-up batch is considered (1 = yes, 0 = no) and then the number of batches.
  - Finally, the sixth row contains the number of arrivals per batch.
- If you want to use files for input ant output, use
  mm1 < in.txt > out.txt
  Of course you can use different file names instead of in.txt or out.txt. Just make sure that the input file exists, as the program does not check for this.
- In Excel, you can copy a worksheet you developed once and refresh the data just by changing the file name. Just make sure that your formulae still stretches the relevant fields after a new import.

Report

In the report, please document

- Objective results: Numerical results in form of tables and/or diagrams
- Subjective results: Your observations, interpretations and comments

exactly in this order! **Do not replicate the theory**; use it rather for understanding and underpinning your observations!

Remember that you cheat yourself if you copied somebody else's report.

Copying will be discovered and prosecuted!

Good luck… and have fun ☺