

Simulation and Performance Evaluation

Homework 4

Blascovich Alessio and Di Noia Matteo

1st June 2025

Simulation Set-Up

In the simulation the times, between two consecutive packet arrivals and the service time for a packet, follow an exponential distribution of parameters λ and μ respectively.

The simulation consists in a FIFO queue of packets arrived that need to be served by a processor.

For measuring it we computed the mean estimators of the parameters searched and also their variance which was used to compute a confidence interval (CI) with 95% confidence level.

Exercise 1

We can start by showing how a simulation evolves over time with particular focus on the number of packets in the queue and in processing. Assuming a single packet can be processed at a time.

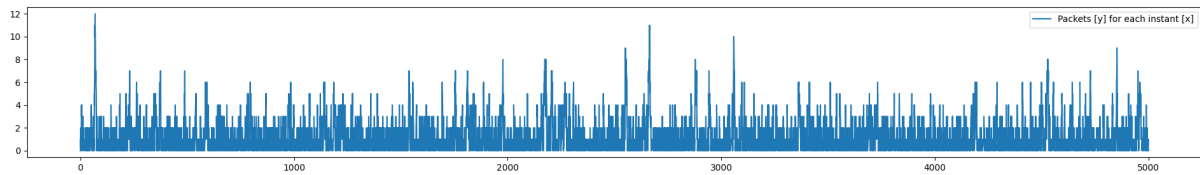


Figure 1: Number of packets in the queue throughout a simulation with $\lambda = 1$ and $\mu = 2$

We started by executing the simulation and plotting the distribution of time in which the simulation had a number of packet, given by summing the queue's elements plus the packet that is being processed, if present. By doing so we found out as expected that such distribution is also an exponential. We can be sure of this by plotting with logarithmic scale, compute the line that best fit the new data and compute the fit of such a line. The resulting line in the case of image ?? is $-0.385 \cdot x + 9.710$. This line is a good fit for our distribution as the R^2 test of fitness of a line returns a value of 0.992 (where 1 is a perfect fit). From this we could easily construct, by elevating to the power of e the line function, the best fit exponential.

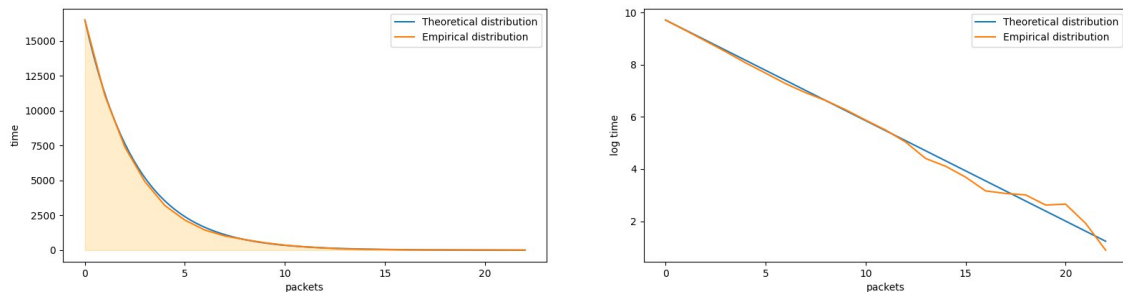


Figure 2: Total length of time in which x packets are in queue or in processing with $\lambda = 1$ and $\mu = 1.5$

Independent Replication

We started using “Independent Replication” as method to gather information about the number of packets in the queue given a time instant. This method has many advantages, especially for finite state simulations, but can be used as well for steady state analysis; although you have to take into consideration

some possible issues such as initialization problem, introducing periodicity and some others. As such we used this method for qualitative analysis and to compare the other methods used.

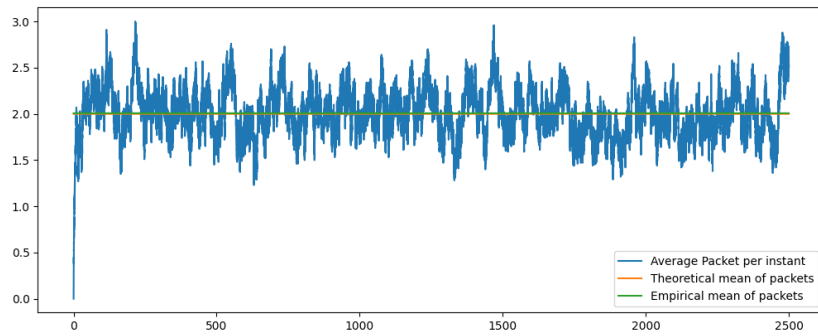


Figure 3: Independent replication with initialization bias, $\lambda = 1$ and $\mu = 1.5$

We started using “Independent Replication” as method to gather information about the number of packets in the queue given a time instant. This method has many advantages, especially for finite state simulations, but can be used as well for steady state analysis; although you have to take into consideration some possible issues such as initialization problem, introducing periodicity and some others. As such we used this method for qualitative analysis and to compare the other methods used.

Using “Independent Replication” we had to take into consideration the initialization bias that is clearly visible in the left hand-side of Figure 3. This bias is caused by all simulations starting with an empty queue, i.e. 0 packets, at the same time namely instant 0. To overcome this problem there are many solutions:

- Starting the simulation at a different initial time, starting gathering data after all the simulation have started;
- Running the simulation for a longer period of time;
- Discard the first data, using some criteria.

We chose the latter method by discarding the first n time instants from each simulation, with n defined as a multiple of the average number of packets in the system in stationary conditions. However this technique works only in non-divergent simulations ($\lambda \geq \mu$) as in those cases the average tends to infinite, this is the case for all the other methods that use the average number of packets in the system. In those cases it is possible to only compute a function describing the climb rate of the average number of packets.

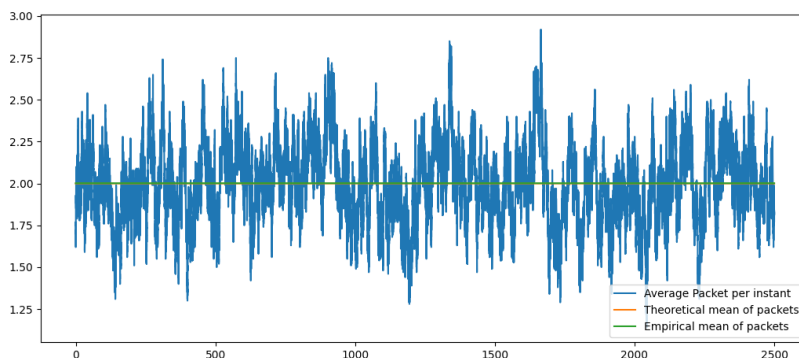


Figure 4: Independent replication without initialization bias, $\lambda = 1$ and $\mu = 1.5$

Discarding the first n time instant we passed from an average of 2.007 ± 0.0438 (Figure 3) to an average of 2.002 ± 0.0405 (Figure 4), with a theoretical value of 2. While the graph in Figure 4 is more representative of the real average over time of a stationary conditions, the obtained average and relative CI do not get much better since the length of the simulation, already partially overwhelm the initialization bias.

Overlapping Batch Means

Then we used “Overlapping Batch Means” as method to estimate data mean and variance. Since for this method only one simulation is involved, we decided to augment the length of it. The new length used was *number of simulations* \times *length of single simulation* used for the “Independent Replication” method.

To use this method we must have:

- Normally distributed batch data;
- IID data.

The second condition is easily verified since every time interval is generated independently and with the same parameters. For the former condition is verified if the batch is large enough. To satisfy both conditions we picked as batch size *simulation length*/1000 and as number of batch 10000 so that they are also overlapping for sure.

Given the theoretical mean is still 2, since we used the same parameters. Some of the empirical means obtained are 1.987 ± 0.0058 , 2.006 ± 0.0059 , 2.010 ± 0.0062 and 1.997 ± 0.0058 . The CIs obtained are strongly under-estimate for the confidence level of 95%, to reach the confidence level the obtained CIs should be doubled in size. This problem is due to the fact that this method tends to under-estimate the variance and the fact that probably we are not completely fulfilling the prerequisites of this method.

To conclude, we expected to achieve better results using this method; however the “Independent Replication” method gave better results (especially confidence intervals) and was easier to work with.

Varying parameters

We then analyzed the convergence of the system with different parameters λ and μ . To be exact we analyzed the system behavior for different value of ρ which is defined as the rate $\frac{\lambda}{\mu}$. We expect value of $\rho \geq 1$ to diverge, value closer to 1 to take accumulate larger queue with higher probability while value closer to 0 to tend to be have smaller queue and dispatch them quickly.

As it is possible to see in the two rows: ?? and ?? the value follow what we expected. In the first ones, the two graph show two simulation with $\rho \geq 1$, and as such diverging. The difference between the two is that while the first having a much more rapid rate of climb, almost constant; the other climb much slower. On the second row instead we see two cases with $\rho \leq 1$, in this case we see that the first example almost resemble the divergent case, creating large queue; instead the second has a much smaller ρ and as such most of packets are quickly dispatched and because of this the queue very often empty with at most a packet currently being processed.

Exercise 2

To compute the variance of the mean two methods where employed: the naive method, which is the variance of all the measurements and the post-stratification, defining the stratum as the number of packet in queue in the system when the packet arrive. The results of this two methods when