

Report HW4

Simone De Carli Damiano Salvaterra
simone.decarli@studenti.unitn.it damiano.salvaterra@studenti.unitn.it

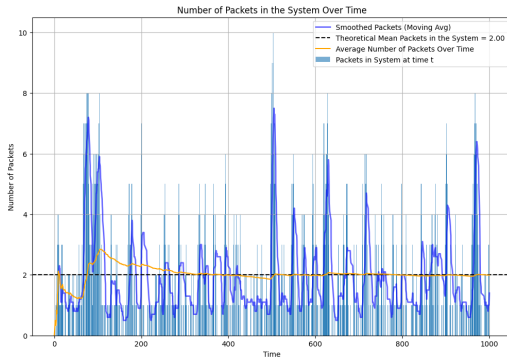
Exercise 1

We implemented a MM1 queue simulator that consists of three main components:

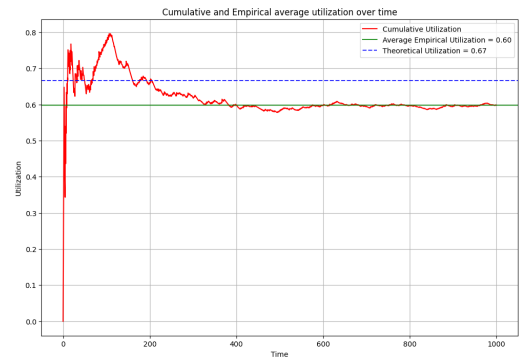
- **Event:** represents the arrival or departure at a certain time;
- **Scheduler:** given a drawn event, it inserts in the min-heap of the future events;
- **MM1:** represents the internal queue system allowing to enqueue and dequeue packets, storing the current number of packets in the queue and if the server is busy or not;
- **Simulator:** act as a controller, it initializes the system and orchestrates the previous components. It also logs to a file the states of the system at each event. Terminates the simulation when a stopping condition is met (time limit, number of events, etc.).

Part 1: Show how the number of packets in the system changes over time

We ran a simulation with $\lambda = 1$ and $\mu = 1.5$ stopping after 1000 time units. Figure 1a shows the number of packets in the system over time and, in yellow, the moving average of the number of packets in the system that after a warm-up period visually converges to the theoretical value of $\frac{\rho}{1-\rho} = 2$. The server utilization is shown in Figure 1b, where the moving average converges to a slightly lower value (0.6) than the theoretical value of $\rho = \frac{\lambda}{\mu} = \frac{2}{3} \approx 0.67$.



(a) Number of packets in the system over time.



(b) Server utilization over time.

Figure 1: The number of packets in the system and the server utilization over time. The empirical data tends to converge to the theoretical values for $\lambda = 1$ and $\mu = 1.5$.

Part 2: Test the simulator with different values of λ and μ

For all combinations of $\lambda \in \{0.5, 1, 1.5\}$ and $\mu \in \{1, 1.5, 2\}$, stopping after 1000 time units. For each combination, we ran 25 independent replications and computed the average empirical utilization.

Figure 2 shows the average empirical utilization over time (averaged through the 25 replications) for each combination of λ and μ . It is possible to see that only blue lines (where $\lambda < \mu$) converge to a stable value, while the other two (red and yellow) saturate to 1, indicating that the system is unstable.

After visualizing the warming up period, we tried to estimate its length visually. We dropped the first 100 time units and computed the average utilization over the remaining 900 time units. Figure 3 shows the truncated results, it is possible to see that the average utilization in the early period is smoother than in previous results.

The number of truncated time units was chosen for all combinations. This was not optimal since it suits better some combinations than others.

In Figure 4 we compare the average utilization before and after truncation only for the stable settings ($\lambda < \mu$). It shows a very small difference between the two, even a slight degradation for $\lambda = 1.0$ and $\mu = 1.5$ where higher frequency probably needed a smaller truncation period.

Also, the impact of the truncation is very small because we used independent replications that decrease the impact of the warm-up period in the first place.

Averaged Cumulative Utilization Over Time

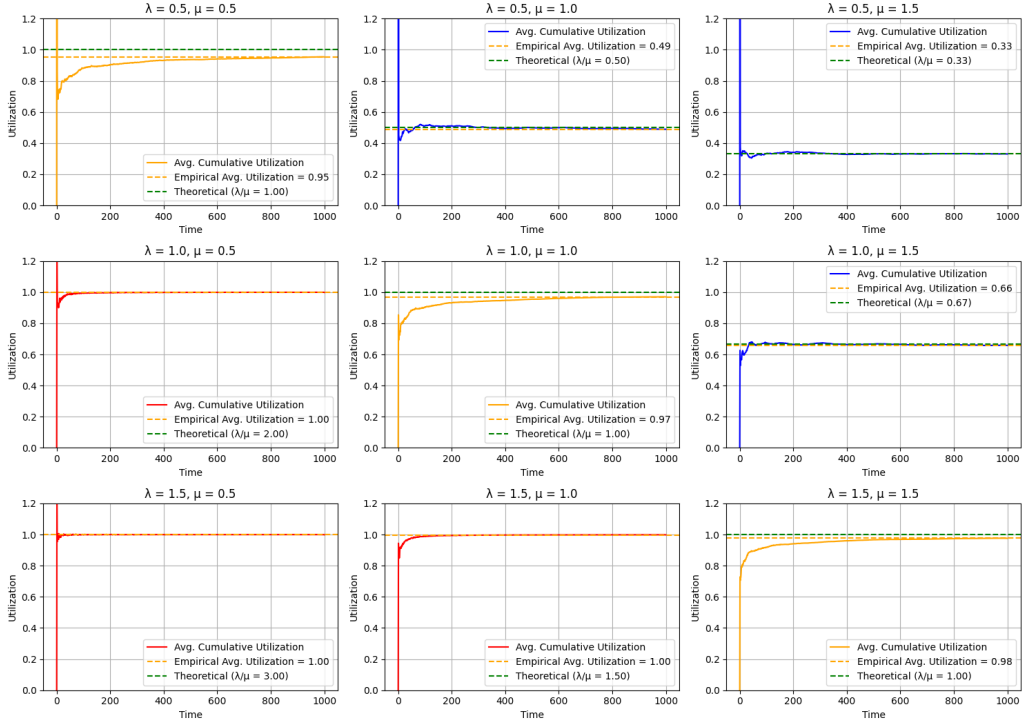


Figure 2: Average empirical utilization for different values of λ and μ . The theoretical value is shown as a dashed line. Red lines are the one with $\lambda > \mu$, yellow $\lambda = \mu$ and blue $\lambda < \mu$ where the system is stable.

Averaged Cumulative Utilization Over Time

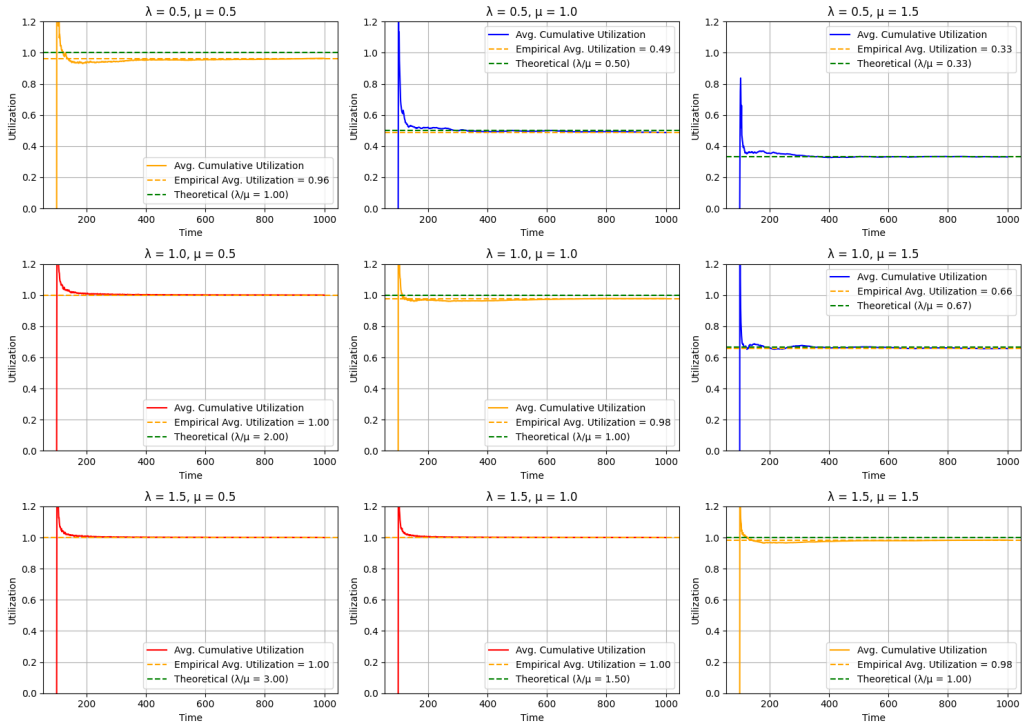


Figure 3: Truncated average empirical utilization for different values of λ and μ .

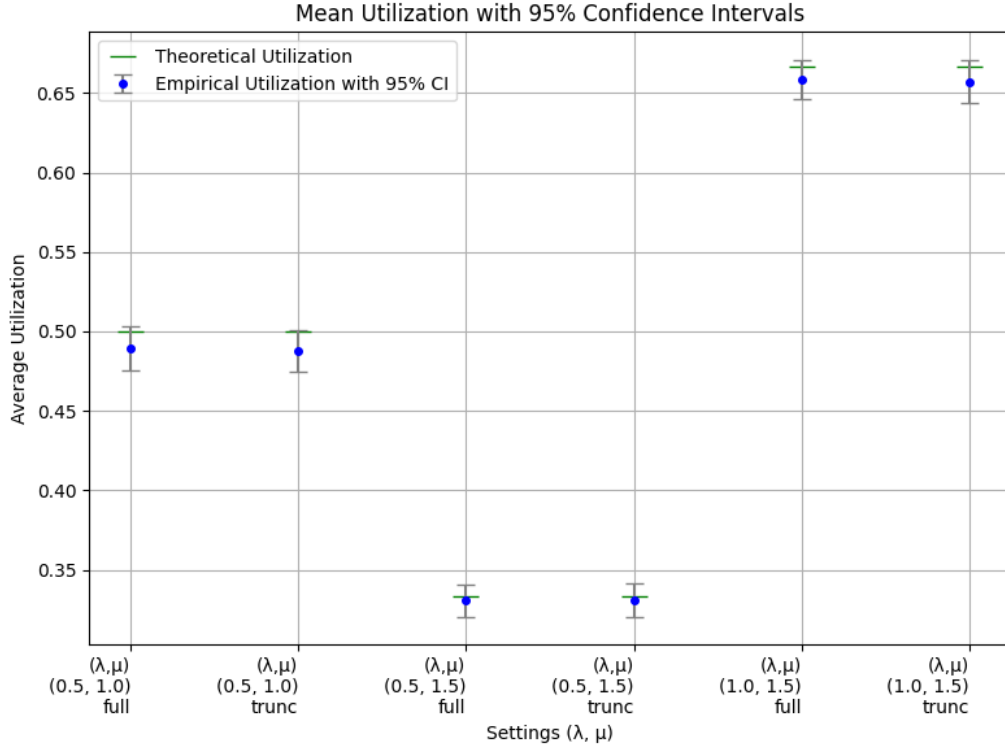


Figure 4: Comparison of the average utilization before and after truncation for the stable settings ($\lambda < \mu$).

Exercise 2

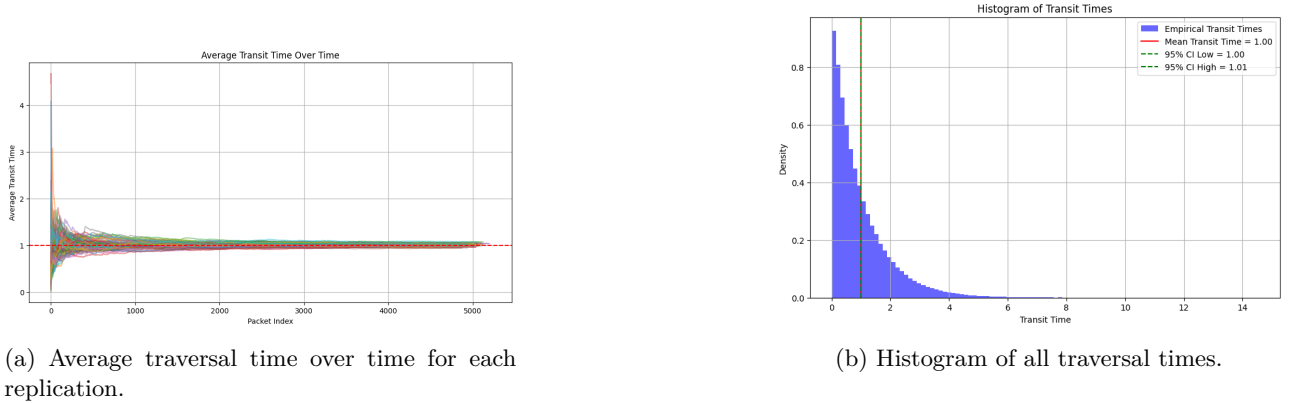
We ran a simulation with:

- $\lambda = 0.5$;
- $\mu = 1.5$;
- 100 independent replications;
- stopping after 10,000 time units.

With these results, we estimated the average traversal time of the packets in the system.

Naive estimator

The naive estimator is simply the average of the traversal times of all packets in the system, computed as the difference between the time of arrival and the time of departure. In Figure 5 we show the results of the naive estimator.



(a) Average traversal time over time for each replication.

(b) Histogram of all traversal times.

Figure 5: Average traversal time over time and histogram of all traversal times for the naive estimator.

Using queue length as control variate

We used the average queue length as control variate to decrease the variance of the estimator. The average queue length is computed as the weighted average of the queue length at each event, where the weight is the time elapsed since the last event.

We computed the linear regression between the average traversal time T and the average queue length L from each replication, obtaining the following model:

$$T = a + b \cdot L + e \quad (1)$$

where a is the intercept, b is the slope and e is the error term.

Then the average traversal time estimator is computed as:

$$\hat{T} = T - b(L - \mu_L) \quad (2)$$

where $\mu_L = \frac{\rho^2}{1-\rho}$ is the theoretical average queue length.

Figure 6 shows the average traversal time and the average queue length for each replication. It is possible to see how the control variate removes the correlation between the two variables, making the average traversal time more stable across replications.

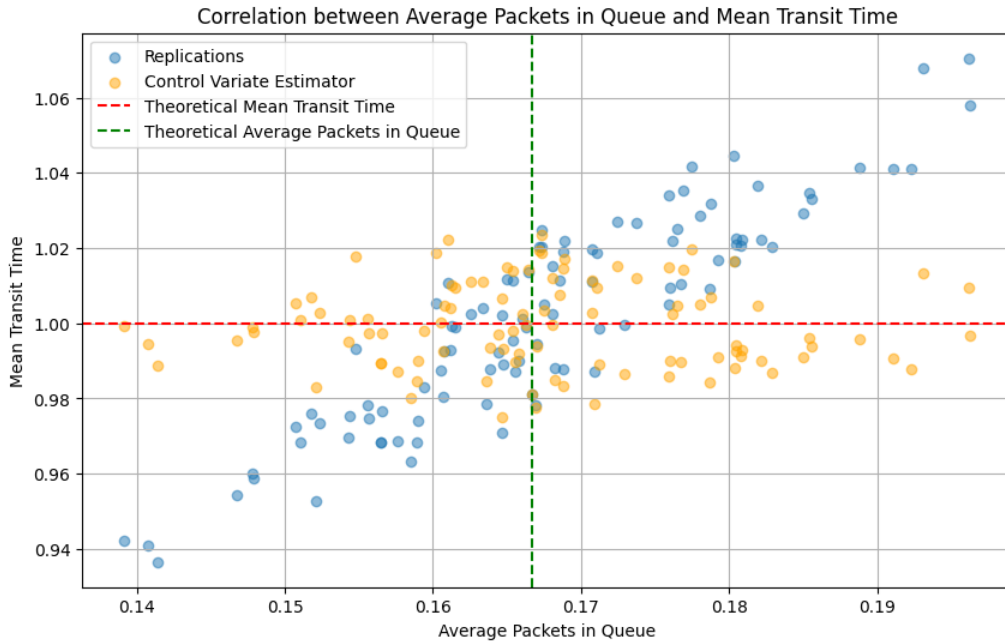


Figure 6: In blue points are the average traversal time and the average queue length for each replication. The orange points are the average traversal time after applying the control variate.

Using post-stratification over packets in the system at the time of arrival

The last estimator we implemented is based on post-stratification, where we group the traversal times by the number of packets in the system at the time of arrival (strata).

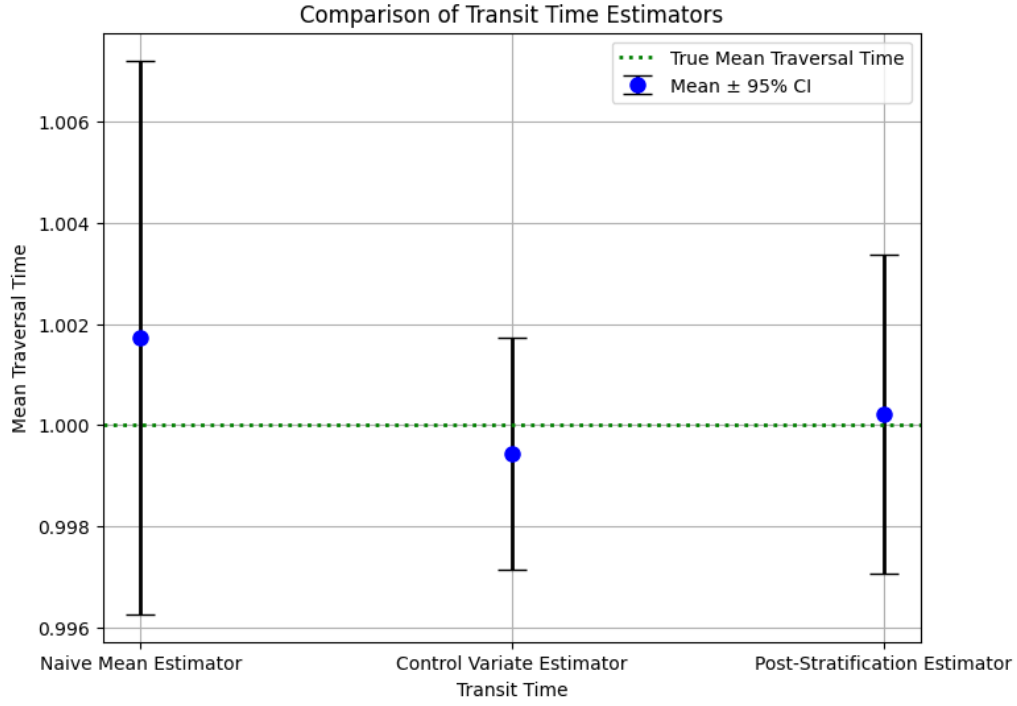
After stratifying the packets, we compute the average traversal time for each stratum and the corresponding probability of being in that stratum: $\rho^k \cdot (1 - \rho)$, where k is the number of packets in the system. Then we compute the weighted average of the traversal times, where the weights are the probabilities of being in each stratum. We must also take into account the tail of the distribution, so for the last stratum k_{\max} we use the probability of being in that stratum or higher, which is:

$$P(X \geq k_{\max}) = \sum_{k=k_{\max}}^{\infty} P(X = k) = \sum_{k=k_{\max}}^{\infty} \rho^k \cdot (1 - \rho) = (1 - \rho) \cdot \rho^{k_{\max}} \cdot \sum_{k=0}^{\infty} \rho^k = (1 - \rho) \cdot \rho^{k_{\max}} \cdot \frac{1}{1 - \rho} = \rho^{k_{\max}} \quad (3)$$

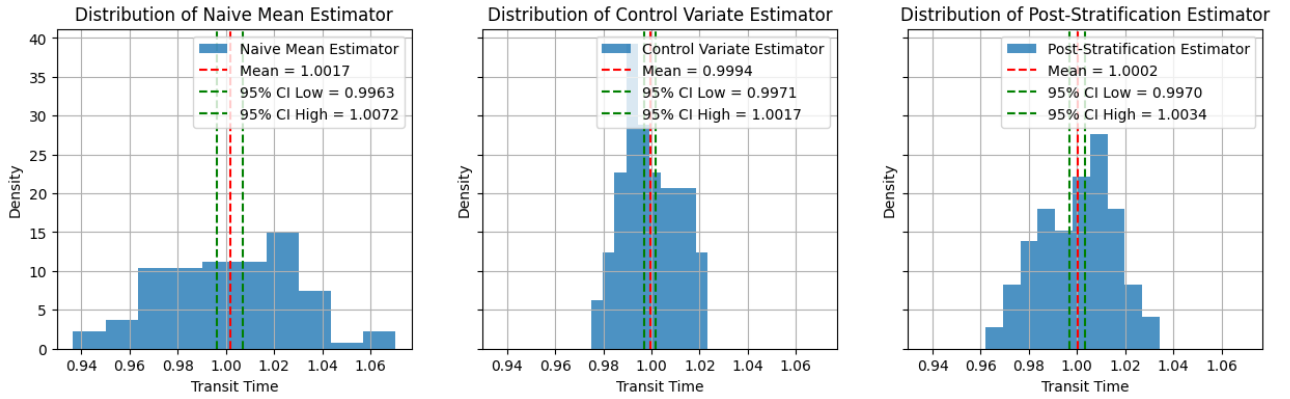
Results

In Figure 7 we show the results of the three estimators. We immediately notice the variance reduction of the in the histogram of the average traversal times (Figure 7b). The naive estimator is obviously the worst one, with

a very high variance in comparison to the other two. The control variate estimator (Figure 7a) has a lower variance than the post-stratification estimator, but the latter provides a more precise estimate.



(a) Average traversal time for the different estimators and 95% confidence intervals.



(b) Histograms of the average traversal times of each replication.

Figure 7: Results of each average traversal time estimator.