Recipe 2: Queueing

Step 0: Write down all the parameters (λ, τ, s, CV_a, CV_s).

Customer parameters

- Arrival rate, λ: number of customer *arriving* per unit time
- Average interarrival time (AIT): average time between two consecutive arrivals, $\text{AIT} = \frac{1}{\lambda}$
- Coefficient of variation for arrivals, CV_a: extent of variability of arrivals in relation to their mean,

  $$CV_a = \frac{\text{standard deviation of interarrival time}}{\text{average interarrival time}}$$

  Example: Poisson arrivals of mean λ have Exponentially distributed interrarival times (std.dev. $=$ mean $= \frac{1}{\lambda}$ for the expenential distriubtion), hence $CV_a = \frac{1/\lambda}{1/\lambda} = 1$.

Service parameters

- Service rate (capacity), μ: number of customers *served* per unit time
- Average service time, τ: average time it takes to serve one customer, $\tau = \frac{1}{\mu}$
- Coefficient of variation for service, CV_s: extent of variability of services offered in relation to their mean,

  $$CV_s = \frac{\text{standard deviation of service time}}{\text{average service time}} = \frac{\text{standard deviation of service time}}{\tau}$$

  Example: Exponential service times: st. dev. equals to the mean, thus here $CV_s = 1$.

Performance measures

- L_q: average number of customer waiting in queue
- L: average number of total customers waiting in the system (queued $+$ on service)
- W_q: average waiting time in the queue
- W: average total waiting time in the system (queue $+$ service), $W = W_q + \tau$
- Little's Law: $L_q = \lambda \cdot W_q$, and $L = \lambda \cdot W$

Step 1: Find the number of servers, s.

Step 2: Find the utilization of a server, ρ.

$$\rho = \frac{\lambda}{\mu \cdot s} = \frac{\lambda \cdot \tau}{s}$$

Step 3: Calculate the average waiting time in the queue.

- Case 1: If $s = 1$ server exists, then

$$W_q = \tau \frac{\rho}{1-\rho} \frac{1}{2} \left(CV_a^2 + CV_s^2\right)$$

- Case 2: If $s > 1$ (there are more than 1 servers) then calculate W_q from the tables. You may use Thales' Law here.

Step 4: Find the average total waiting time in the system.

Average total waiting time in the system (queue + service):

$$W = W_q + \tau$$

Step 5: Find the average number of customers.

Use Little's Law: the average queue lenght is

$$L_q = \lambda \cdot W_q,$$

and the average total number of customers in the system is

$$L = \lambda \cdot W$$