Preventing queues growing large

QUEUEING THEORY

Youri Raaijmakers

8 March, 2021

Department of Mathematics and Computer Science

Queueing theory?

In Google search: 1.940.000 results (in 0.61 seconds)

Wikipedia: Queueing theory is the mathematical study of waiting lines, or queues.

Queueing – word with the second most consecutive vowels (five)

TU/e

Queueing theory?

Waiting in line at the supermarket

Queueing theory?

Waiting in line at the toll highway

Queueing theory?

(Virtually) waiting in wireless communication networks

What is there to study?

Which waiting line should I take?

How many check outs are needed?

How can we reduce the waiting times?

Why is the other line always going faster?

What is there to study?

(Customer perspective) - Which waiting line should I take?

(Design perspective) - How many check outs are needed?

(Operator perspective) - How can we reduce the waiting times?

(Philosophical perspective) - Why is the other line always going faster?

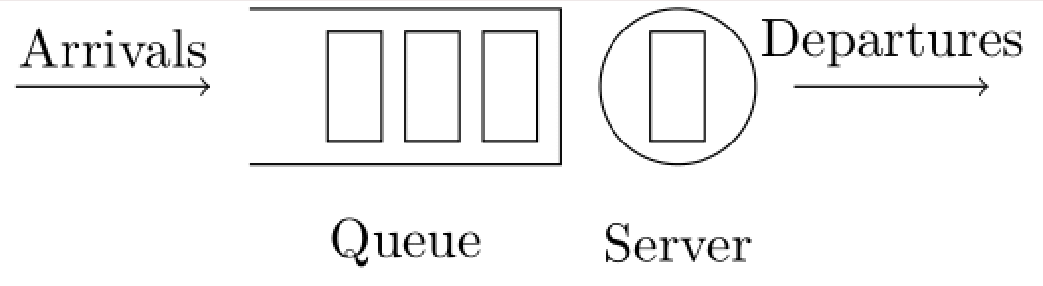
Queueing theory – Mathematical model

Explain reality as good as possible
- Complexity
- Slow solutions
- Accurate
- Detailed

Ability to analyse the mathematical model
- Simplicity
- Fast solutions
- Approximate
- Concise

Mathematics

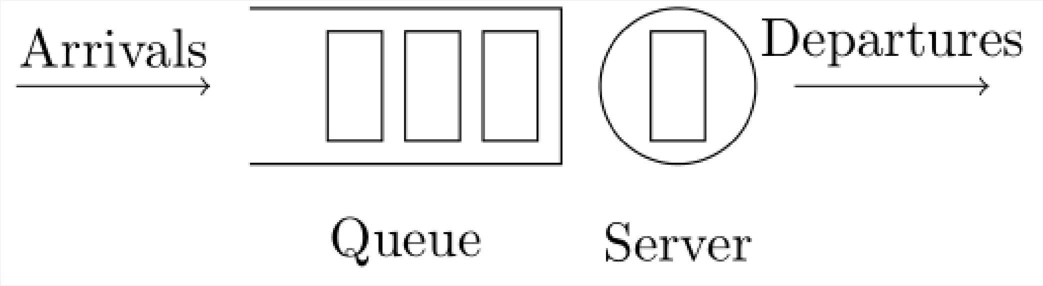
M/M/1



L denotes the number of customers in the supermarket
Goal: Calculate $\mathbb{P}(L = i)$, the probability that there are i customers in the supermarket

Mathematics

M/M/1

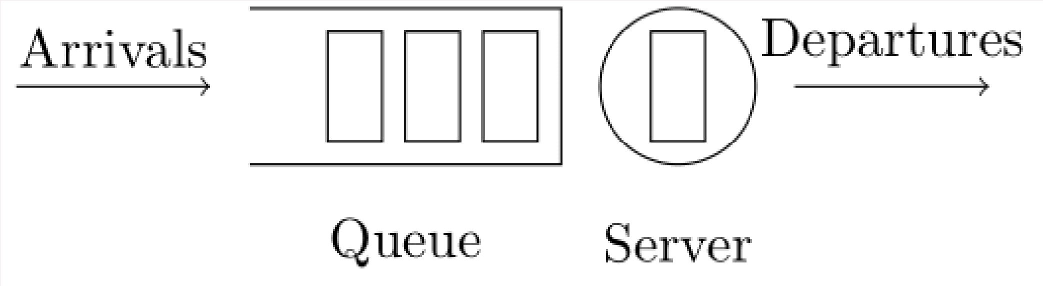


What can happen when there is 1 customer in the supermarket?

L: total customers
λ: arrivals
μ: departures

Mathematics

M/M/1

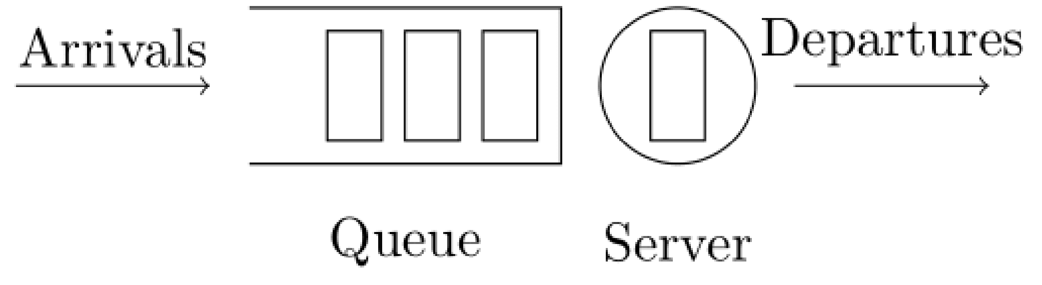


What can happen when there are i customers in the supermarket?

L: total customers
λ: arrivals
μ: departures

Mathematics

M/M/1



Goal: Calculate $p_i = \mathbb{P}(L = i)$

$\lambda p_0 = \mu p_1 \quad \rightarrow p_1 = p_0 \lambda/\mu$

$\lambda p_1 = \mu p_2 \quad \rightarrow p_2 = p_1(\lambda/\mu) = p_0(\lambda/\mu)^2$

$\lambda p_2 = \mu p_3 \quad \rightarrow p_3 = p_2(\lambda/\mu) = p_0(\lambda/\mu)^3$

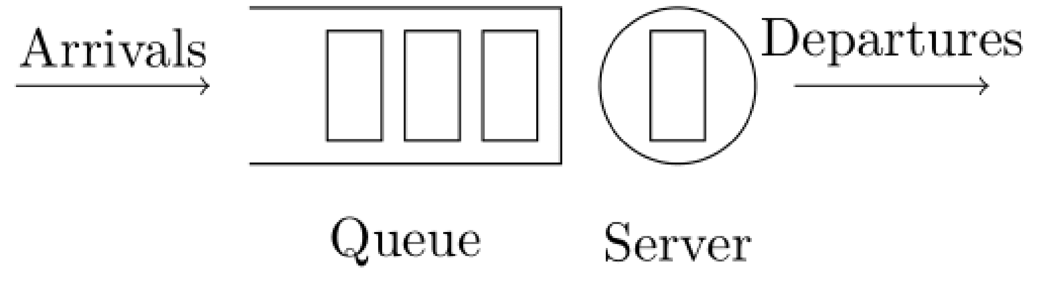
….

L: total customers
λ: arrivals
μ: departures

TU/e

Mathematics

M/M/1



Goal: Calculate $p_i = \mathbb{P}(L = i)$

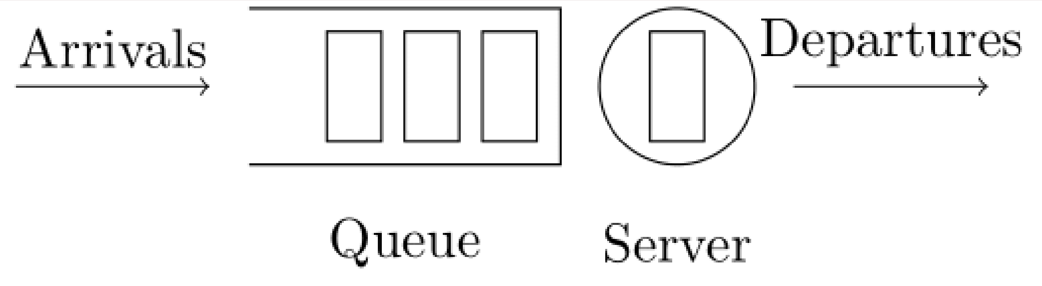
… calculations … $p_i = \left(\frac{\lambda}{\mu}\right)^i \left(1 - \frac{\lambda}{\mu}\right)$

L: total customers
λ: arrival rate
μ: departure rate

TU/e

Mathematics

M/M/1



$$\mathbb{P}(L = i) = p_i = \left(\frac{\lambda}{\mu}\right)^i \left(1 - \frac{\lambda}{\mu}\right)$$

The expected number of customers in the supermarket is $\mathbb{E}[L] = \frac{\rho}{1-\rho}$ with $\rho = \frac{\lambda}{\mu}$

Question M/M/1

Think of a supermarket in your area. We model the checkout of this supermarket as an M/M/1 queue with an arrival rate of 60 customers per hour, thus $\lambda = 1$, and a mean service time of 45 seconds, thus $\mu = 1.33$.

a) Draw the Flow diagram for this M/M/1 queue
b) Determine the probabilities $\mathbb{P}(L = 0)$, $\mathbb{P}(L = 1)$ and $\mathbb{P}(L = 2)$

We call the supermarket crowded when there are 3 or more customers present

c) On average, what is the fraction of the day that this supermarket is crowded

Solutions M/M/1

Draw the Flow diagram for this M/M/1 queue

a) 

Solutions M/M/1

Determine the probabilities $\mathbb{P}(L = 0)$, $\mathbb{P}(L = 1)$ and $\mathbb{P}(L = 2)$

b) $\mathbb{P}(\mathrm{L} = 0) = p_0 = \rho^0(1 - \rho)^1 = (1 - \rho) = \left(1 - \frac{1}{1.33}\right) = 0.25$

$\mathbb{P}(\mathrm{L} = 1) = p_1 = \rho^1(1 - \rho)^1 = \frac{1}{1.33}\left(1 - \frac{1}{1.33}\right) = 0.187$

$\mathbb{P}(\mathrm{L} = 2) = p_2 = \rho^2(1 - \rho)^1 = \left(\frac{1}{1.33}\right)^2\left(1 - \frac{1}{1.33}\right) = 0.140$

Solutions M/M/1

We call the supermarket crowded when there are 3 or more customers present

On average, what is the fraction of the day that this supermarket is crowded

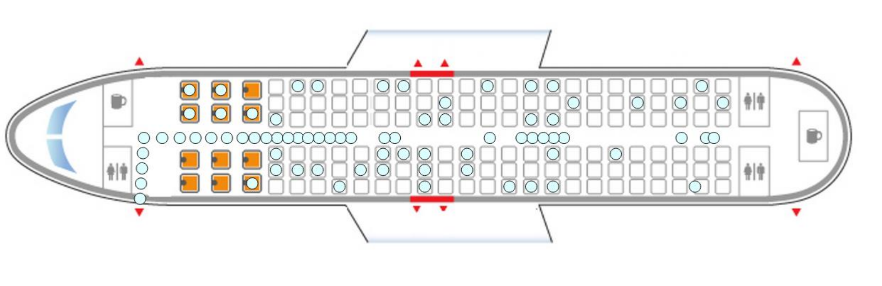
c) $\mathbb{P}(\mathrm{L} \geq 3) = 1 - \mathbb{P}(\mathrm{L} = 0) - \mathbb{P}(\mathrm{L} = 1) - \mathbb{P}(\mathrm{L} = 2)$

$\mathbb{P}(\mathrm{L} \geq 3) = 1 - 0.25 - 0.187 - 0.140 = 0.423$

Airport boarding

When boarding the plane some passengers try to be the first at the line, but have you ever wondered what would be the best strategy to board?

1) Random
2) Rear first
3) Front first
4) Windows/middle/aisle, random
5) Windows/middle/aisle, ordered

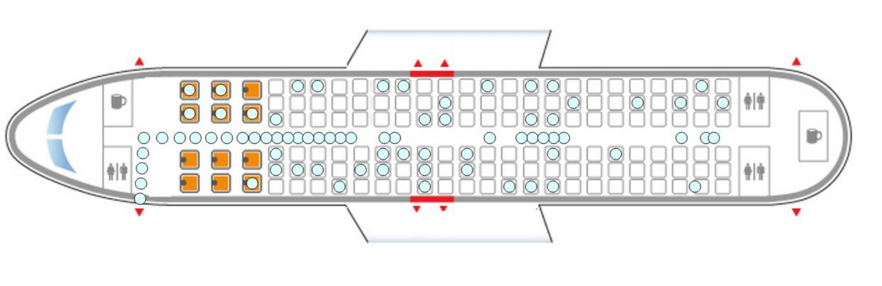


Airport boarding

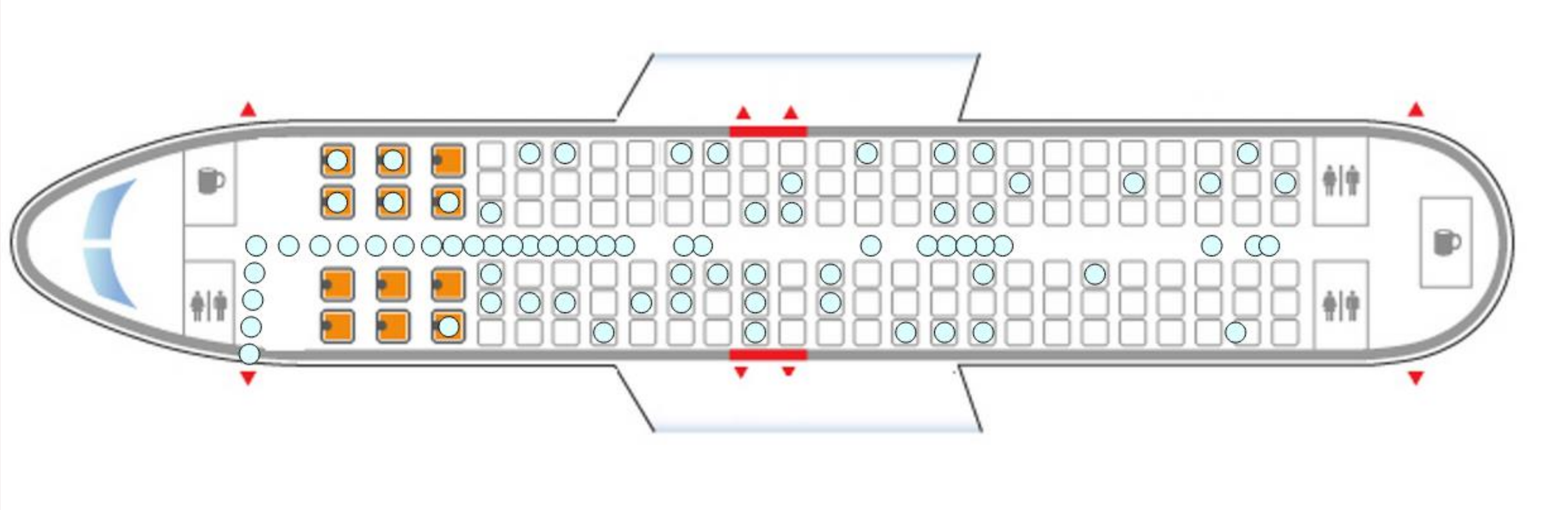
When boarding the plane some passengers try to be the first at the line, but have you ever wondered what would be the best strategy to board?

Say 1) random takes 10 min
How long does 5) Windows/middle/aisle, ordered take?

a) 10 min
b) 7.5 min
c) 5 min
d) 2.5 min



Airport boarding

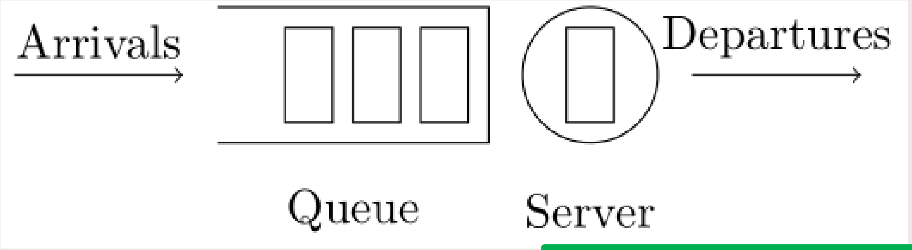


https://www.networkpages.nl/CustomMedia/Animations/Queueing/AirportBoarding.html

Mean value approach

- Analyze queueing systems with general service times M/G/1
- One expression we derive for the expected waiting time

Mean value approach

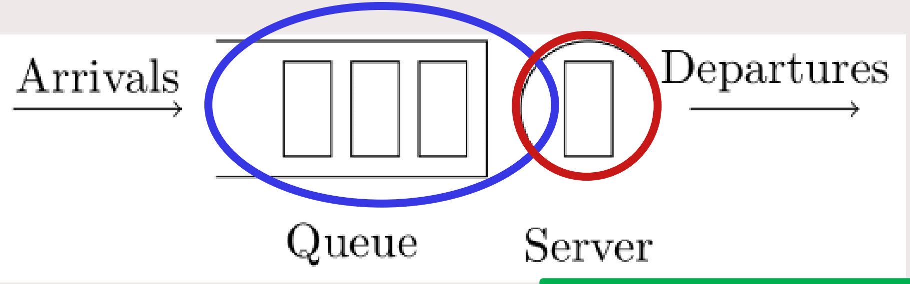


$$\mathbb{E}[W] = \mathbb{E}[L^Q]\mathbb{E}[B] + \rho\mathbb{E}[R]$$

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation

TU/e

Mean value approach



$$\mathbb{E}[W] = \mathbb{E}[L^Q]\mathbb{E}[B] + \rho\mathbb{E}[R]$$

- $\mathbb{E}[L^Q]\mathbb{E}[B]$ the expected time the arriving customer has to wait for customers in the queue
- $\rho\mathbb{E}[R]$ the expected time the arriving customer has to wait for the customer in service

  $\rho = \mathbb{P}(L > 0)$

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation

Little’s law

$$\mathbb{E}[L] = \lambda\mathbb{E}[S]$$

L: total customers
λ: average number customers entering
S: sojourn time

All customers pay 1 euro per unit time while in the system

- Let customers pay continuously in time.
  Average reward is $\mathbb{E}[L]$ euro per unit time
- Let customers pay for their residence when they leave.
  Customers stay on average $\mathbb{E}[S]$ time units in the system
  Average number of customers leaving is λ since everyone that enters should also leave the system

Little's law

$$\mathbb{E}[L^Q] = \lambda \mathbb{E}[W]$$

L: total customers
λ: average number customers entering
S: sojourn time

All customers pay 1 euro per unit time while in the queue

- Let customers pay continuously in time.
  Average reward is $\mathbb{E}[L^Q]$ euro per unit time
- Let customers pay for their residence when they leave.
  Customers stay on average $\mathbb{E}[W]$ time units in the queue
  Average number of customers leaving is λ since everyone that enters should also leave the system

Mean value approach

$$\mathbb{E}[W] = \mathbb{E}[L^Q]\mathbb{E}[B] + \rho\mathbb{E}[R]$$
$$\mathbb{E}[L^Q] = \lambda\mathbb{E}[W]$$

….

$$\mathbb{E}[W] = \frac{\rho\mathbb{E}[R]}{1-\rho}$$

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation
$\rho = \lambda\mathbb{E}[B]$

Mean value approach

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation
$\rho = \lambda \mathbb{E}[B]$

$$\mathbb{E}[W] = \mathbb{E}[L^Q]\mathbb{E}[B] + \rho\mathbb{E}[R]$$
$$\mathbb{E}[L^Q] = \lambda\mathbb{E}[W]$$

$\rightarrow \mathbb{E}[W] = \lambda\mathbb{E}[W]\mathbb{E}[B] + \rho\mathbb{E}[R]$
$\rightarrow \mathbb{E}[W] = \rho\mathbb{E}[W] + \rho\mathbb{E}[R]$
$\rightarrow \mathbb{E}[W] - \rho\mathbb{E}[W] = \rho\mathbb{E}[R]$
$\rightarrow (1 - \rho)\mathbb{E}[W] = \rho\mathbb{E}[R]$

$$\mathbb{E}[W] = \frac{\rho\mathbb{E}[R]}{1 - \rho}$$

Conservation law

So far we considered the first-come first-served (FCFS) discipline, but there are other disciplines:

- Random order
- Last-come first-served (LCFS)
- Priorities

Waiting time depends on the discipline

Conservation law

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation

Consider r types of customers:

- Each type arrives with rate λ_i
- Expected service time $\mathbb{E}[B_i]$
- Expected residual service time $\mathbb{E}[R_i]$

Next, we derive a conservation law for the expected waiting times of the r type of customers
Weighted sum of these expected waiting times is independent of the service discipline

Conservation law

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation

For service discipline P:

$$\mathbb{E}[V(P)] = \sum_{i=1}^{r} \mathbb{E}\left[L_i^Q(P)\right]\mathbb{E}[B_i(P)] + \rho_i(P)\mathbb{E}[R_i(P)]$$

$\mathbb{E}[V(P)]$ is the expected amount of work in the system

Conservation law

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation

For service discipline P:

$$\mathbb{E}[V(P)] = \sum_{i=1}^{r} \mathbb{E}\left[L_i^Q(P)\right]\mathbb{E}[B_i(\cancel{P})] + \rho_i(\cancel{P})\mathbb{E}[R_i(\cancel{P})]$$

$\mathbb{E}[V(P)]$ is the expected amount of work in the system

Conservation law

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation

For service discipline P:

$$\mathbb{E}[V(\cancel{P})] = \sum_{i=1}^{r} \mathbb{E}\left[L_i^Q(P)\right]\mathbb{E}[B_i] + \rho_i \mathbb{E}[R_i]$$

$\mathbb{E}[V(P)]$ is the expected amount of work in the system
The amount of work decreases with one unit per time
independent of the customer being served

Conservation law

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation
V: workload

For service discipline P:

$$\mathbb{E}[V] = \sum_{i=1}^{r} \mathbb{E}\left[L_i^Q(P)\right]\mathbb{E}[B_i] + \rho_i \mathbb{E}[R_i]$$

$$\mathbb{E}\left[L_i^Q(P)\right] = \lambda_i \mathbb{E}[W_i(P)]$$ (Little's law)

$$\rightarrow \sum_{i=1}^{r} \rho_i \mathbb{E}[W_i(P)] = \mathbb{E}[V] - \rho_i \mathbb{E}[R_i]$$

Conservation law

For service discipline P:

$$\sum_{i=1}^{r} \rho_i \mathbb{E}[W_i(P)] = \mathbb{E}[V] - \rho_i \mathbb{E}[R_i]$$

Improving the expected waiting time of one customer type will always degrade the expected waiting time of another customer type

W: waiting time
L^Q: queue length
B: service time
R: residual service
ρ: load/occupation
V: workload

Than queue