Course: Applied Probability

Chapter: [2]

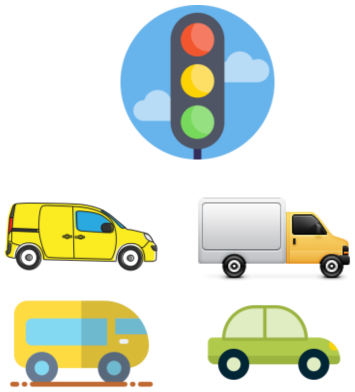
Queuing Theory

Section: [2.1]

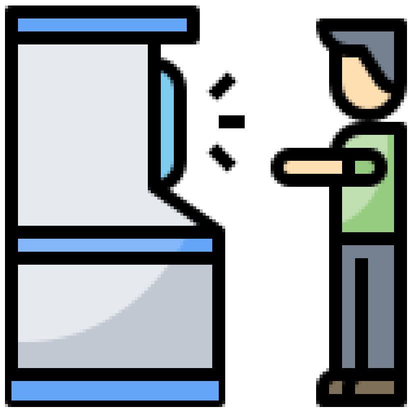
Some Queuing Terminology

Queuing System

Examples Each of us has spent a great deal of time waiting in lines.



Traffic Lights



ATM Banks

Calling Centers

Supermarket Cashier

Description To describe a queuing system, an "**input process**" and an "**output process**" must be specified.

The Input or Arrival Process

Concepts

The input process is usually called the "**arrival process**". Arrivals are called "**customers**".

If more than one arrival can occur at a given instant, we say that "**bulk arrivals**" are allowed.

If a customer arrives but fails to enter the system, we say that the customer has "**balked**".

Models in which arrivals are drawn from a small population are called "**finite source models**".

The Output or Service Process

Concepts

To describe the output process of a queuing system, we usually specify a probability distribution — **the service time distribution** — which governs a customer's service time.

Types of Servers

Servers in Parallel

Servers are in parallel if all server provide the same type of service and a customer need only pass through one server to complete service.

Servers in Series

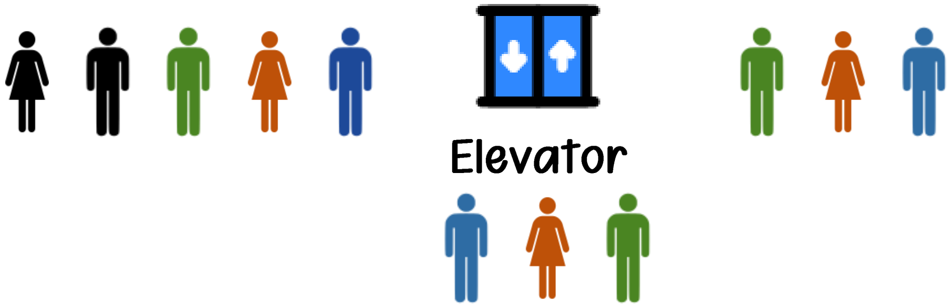
Servers are in series if a customer must pass through several servers before completing service.

Queue Discipline

Concept The queue discipline describes the method used to determine the order in which customers are served.

FIFO The most common queue discipline is the **"first in, first out"**, in which customers are served in the order of their arrival.

LIFO Under the **"last in, first out"**, the most recent arrivals are the first to enter service.



Queue Discipline

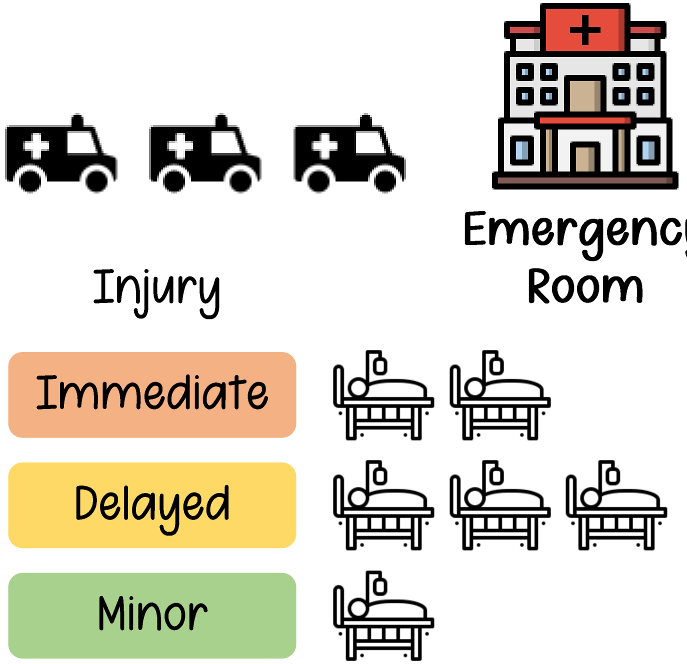
SIRO

If the next customer to enter service is randomly chosen from those customers waiting for service it is referred to as the "**service in random order**".

Ramadan television competitions, where every day a question is asked to the audience, and at the end of the month, a candidate is chosen at random to win the prize.

Priority Discipline

A priority discipline classifies each arrival into one of several "**categories**". Each category is then given a priority level, and within each priority level, customers enter service on an FIFO basis.



Queuing Theory

Important Questions Queuing theory can help us address various questions in a queuing system.

1) What fraction of the time is each server idle?
2) What is the expected number of customers present in the queue?
3) What is the expected time that a customer spends in the queue?
4) What is the probability distribution of the number of customers present in the queue?
5) What is the probability distribution of a customer's waiting time?

⋮

Course: Applied Probability

Chapter: [2]
Queuing Theory

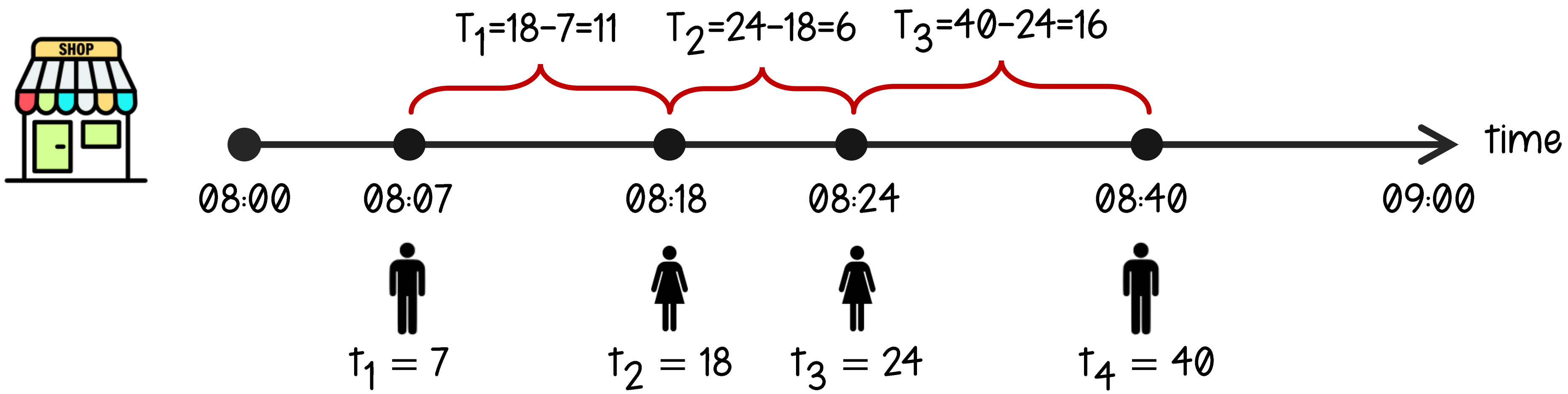
Section: [2.2]
Modeling Arrival and Service Processes

Modeling the Arrival Process

Assumption We assume that at most one arrival can occur at a given instant of time.

We define t_i to be the time at which the i^{th} customer arrives.

For $i \geq 1$, we define $T_i = t_{i+1} - t_i$ to be the i^{th} **interarrival time**.



Modeling the Arrival Process

Definition

We define λ to be the "**average arrival rate**", which will have units of arrivals per hour.

We define $1/\lambda$ to be the "**average interarrival time**" (will have units of hours per arrival).

Example

Determine the average arrival rate per hour if one arrival occurs every 20 minutes.

$$\text{average arrival rate } (\lambda) = \frac{60}{20} \times 1 = 3 \text{ arrivals/hour}$$

$$\text{average interarrival time } (1/\lambda) = \frac{1}{3} \text{ hour/arrival}$$

Modeling the Arrival Process

Example Determine the average interarrival time in hours if the number of arrivals in a 30-minute period is 10.

$$\text{average arrival rate } (\lambda) = \frac{60}{30} \times 10 = 20 \text{ arrivals/hour}$$

$$\text{average interarrival time } (1/\lambda) = \frac{1}{20} \text{ hour/arrival}$$

Note In modeling the arrival process, we assume that the T_i's are **independent**, **continuous** random variables described by the random variable A.

So, the random variable A has a density function $f(t)$.

Modeling the Arrival Process

Distribution of "A"

The most common choice for a distribution of A is the **exponential distribution.**

An exponential distribution with parameter λ has a density $f(t) = \lambda e^{-\lambda t}$; where λ to be the average arrival rate.

Probabilities

$$P(A \leq c) = \int_{0}^{c} \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda c} \quad ; \; c \geq 0$$

negative interarrival time is impossible

$$P(A > c) = 1 - P(A \leq c) = e^{-\lambda c} \quad ; \; c \geq 0$$

Modeling the Arrival Process

Expected Value and Variance

$$E(A) = \int_0^\infty t\,\lambda e^{-\lambda t} dt = \frac{1}{\lambda}$$

average interarrival time

$$Var(A) = \underbrace{\left[\int_0^\infty t^2\,\lambda e^{-\lambda t} dt\right]}_{E\left(A^2\right)} - [E(A)]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Modeling the Arrival Process

Stationary Interarrival Time

An arrival process is said to be **"stationary"** if, for any time interval (e.g., an hour), the average number of arrivals in this time interval only depends on the length of the time interval, not on the starting time of the interval.

No-Memory Property

If A has an exponential distribution, then for all nonnegative values of t and h,

$$P(A > t + h \mid A > t) = P(A > h)$$

$$\begin{aligned} P(A > 9 \mid A > 5) &= P(A > 7 \mid A > 3) \\ &= P(A > 18 \mid A > 14) \\ &= P(A > 4 \mid A > 0) = P(A > 4) = e^{-4\lambda} \end{aligned}$$

Modeling the Arrival Process

Example The time between arrivals at a bank is exponentially distributed with average value 0.04 hour. The bank opens at 8:00 AM.

a) Write the exponential distribution that describes the interarrival time.

$$\frac{1}{\lambda} = .04 \qquad \lambda = 25 \qquad f(t) = 25e^{-25t} \;;\; t > 0$$

b) Find the probability that no customers will arrive at the bank by 8:15 AM.

$$P\left(t > \frac{15}{60}\right) = P(t > 0.25) = e^{-25\times 0.25} \approx 0.00193$$

Modeling the Arrival Process

Example The time between arrivals at a bank is exponentially distributed with average value 0.04 hour. The bank opens at 8:00 AM.

c) It is now 8:35 A.M. The last customer entered the bank at 8:26. What is the probability that the next customer will arrive before 8:38 A.M.?

$$P\left(t < \frac{3}{60}\right) = P(t < 0.05) = 1 - e^{-25\times 0.05} \approx 0.713$$

d) What is the average number of arriving customers between 8:10 and 8:45 A.M?

$$25 \times \frac{45-10}{60} = 25 \times \frac{35}{60} \approx 14.58 \text{ arrivals}$$

Modeling the Arrival Process

Poisson Distribution

A discrete random variable N has a **Poisson distribution** with parameter λ if, for $n = 0,1,2,\cdots$,

$$P(N = n) = \frac{e^{-\lambda}\,\lambda^{n}}{n!}$$

If N is a Poisson random variable, it can be shown that

$$E(N) = Var(N) = \lambda$$

Theorem

Interarrival times are exponential with parameter λ if and only if the number of arrivals to occur in an interval of length t follows a Poisson distribution with parameter λt.

Modeling the Arrival Process

Note If we define N_t to be the number of arrivals to occur during any time interval of length t, then

$$P(N_t = n) = \frac{e^{-(\lambda t)}\,(\lambda t)^n}{n!} \quad \text{for} \quad n = 0,1,2,\cdots$$

If N_t is a Poisson random variable with parameter λt, then

$$E(N_t) = Var(N_t) = \lambda t$$

Theorem If

- the arrival rate is stationary,
- bulk arrivals cannot occur,
- past arrivals do not affect future arrivals,

then N_t follows Poisson with parameter λt, and interarrival times are exponential with parameter λ.

Modeling the Arrival Process

Example The number of cups of coffee ordered per hour at a coffeeshop follows a Poisson distribution, with an average of 30 cups per hour being ordered.

a) Find the probability that exactly 50 cups are ordered between 10 A.M. and 12 midday.

$$\lambda t = (30)(12-10) = 60$$

$$P(N_2 = 50) = \frac{e^{-60} \times 60^{50}}{50!} \approx 0.0233$$

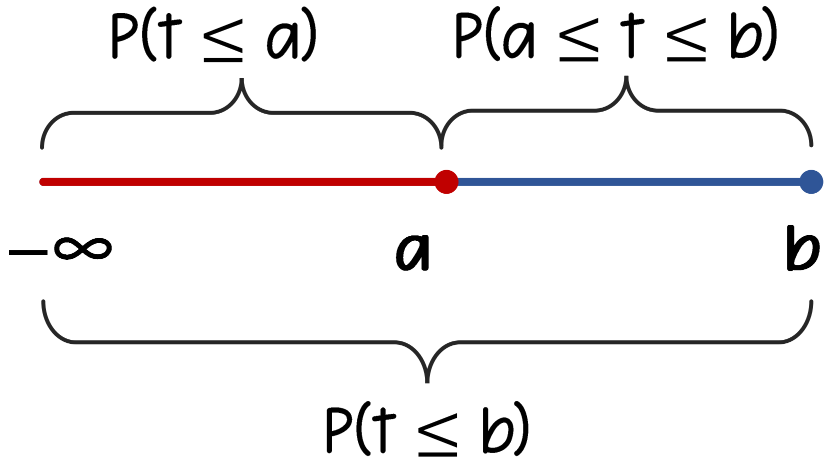
b) Find the mean and variance of the number of coffee cups ordered between 9 A.M. and 1 P.M.

$$E(N_4) = Var(N_4) = (30)(4) = 120$$

Modeling the Arrival Process

Example The number of cups of coffee ordered per hour at a coffeeshop follows a Poisson distribution, with an average of 30 cups per hour being ordered.

c) Find the probability that the time between two consecutive orders is between 1 and 3 minutes.



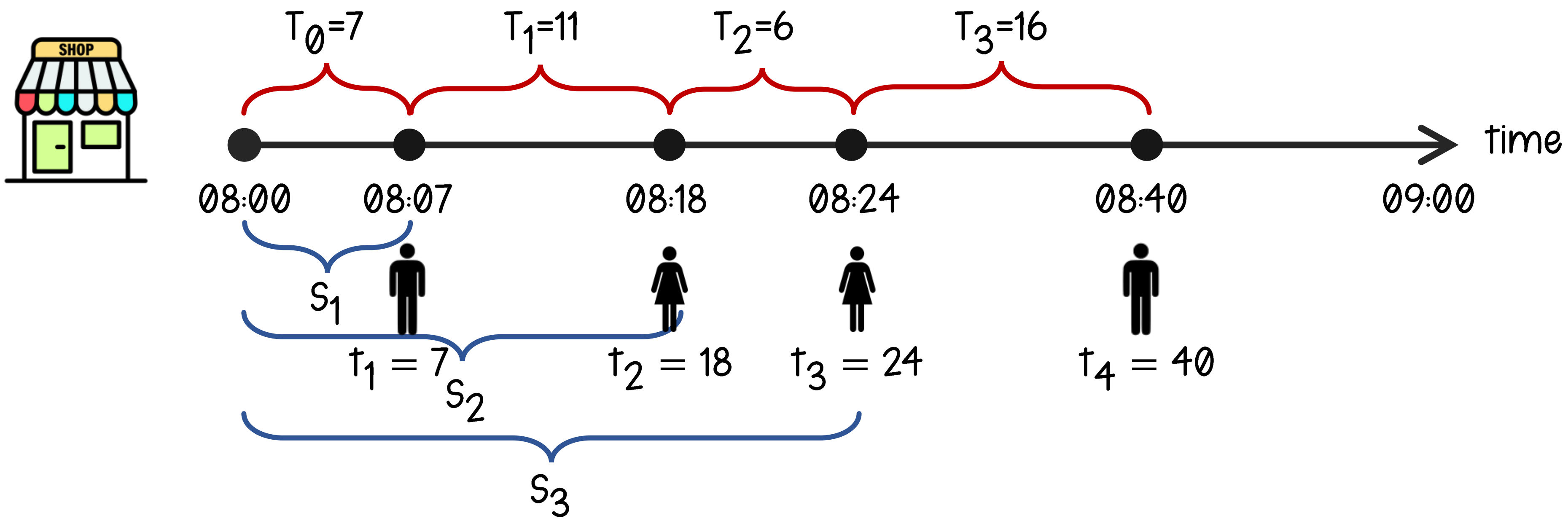
$$P\left(\frac{1}{60} < t < \frac{3}{60}\right) = P\left(t < \frac{3}{60}\right) - P\left(t < \frac{1}{60}\right)$$

$$= \left[1 - e^{-(30)(3/60)}\right] - \left[1 - e^{-(30)(1/60)}\right]$$

$$= e^{-0.5} - e^{-1.5} \approx 0.3834$$

Modeling the Arrival Process

Erlang Distribution



Define the continuous random variable S_k to be the arrival time of customer k.

$$
\begin{aligned}
S_1 &= 7 &&= T_0 \\
S_2 &= 18 &&= T_0 + T_1 \\
S_3 &= 24 &&= T_0 + T_1 + T_2
\end{aligned}
$$

$$\therefore\ S_k = \sum_{j=0}^{k-1} T_j$$

Modeling the Arrival Process

Erlang Distribution

An **Erlang distribution** is a continuous random variable whose density function is

$$f(t)=\frac{\lambda\,(\lambda t)^{\lambda-1}e^{-\lambda t}}{(k-1)!}$$

where λ is the average arrival rate.

If S_k is an Erlang random variable with parameter λ, then

$$E(S_k)=\frac{k}{\lambda} \quad \text{and} \quad Var(S_k)=\frac{k}{\lambda^2}$$

Note

If interarrival times do not appear to be exponential they are often modeled by an Erlang distribution.

Modeling the Service Process

Assumptions

We assume that the service times of different customers are independent random variables.

We call μ the ***average service rate*** that has units of customers per hour.

So, $1/\mu$ is the ***average service time*** that has units of hours per customer.

Example

If $\mu=5$, then if customers were always present, the server could serve an average of 5 customers per hour, and the average service time of each customer would be 1/5 hour = 12 minutes.

Modeling the Service Process

Note In certain situations, interarrival or service times may be modeled as having zero variance; in this case, interarrival or service times are considered to be **deterministic**.

If interarrival times are deterministic, then each interarrival time will be exactly $1/\lambda$, and if service times are deterministic, each customers service time is exactly $1/\mu$.

The Kendall—Lee Notation for Queuing Systems

Interarrival times distribution / Service times distribution / Number of parallel servers / Queue discipline / System Capacity / Population Size

Interarrival times distribution:

M = Exponential
D = Deterministic
E = Erlang
G = General

Service times distribution:

M = Exponential
D = Deterministic
E = Erlang
G = General

Number of parallel servers:

1
2
⋮

Queue discipline:

FIFO
LIFO
SIRO
GD

System Capacity:

1
2
⋮

Population Size:

1
2
⋮

The Kendall—Lee Notation for Queuing Systems

Example M/M/8/FCFS/10/∞

The given queuing system might represent a health clinic with

- 8 doctors,
- exponential interarrival and service times,
- an FCFS queue discipline,
- and a total capacity of 10 patients

Note In many important models 4/5/6 is GD/∞/∞. If this is the case, then 4/5/6 is often omitted

Course: Applied Probability

Chapter: [2]

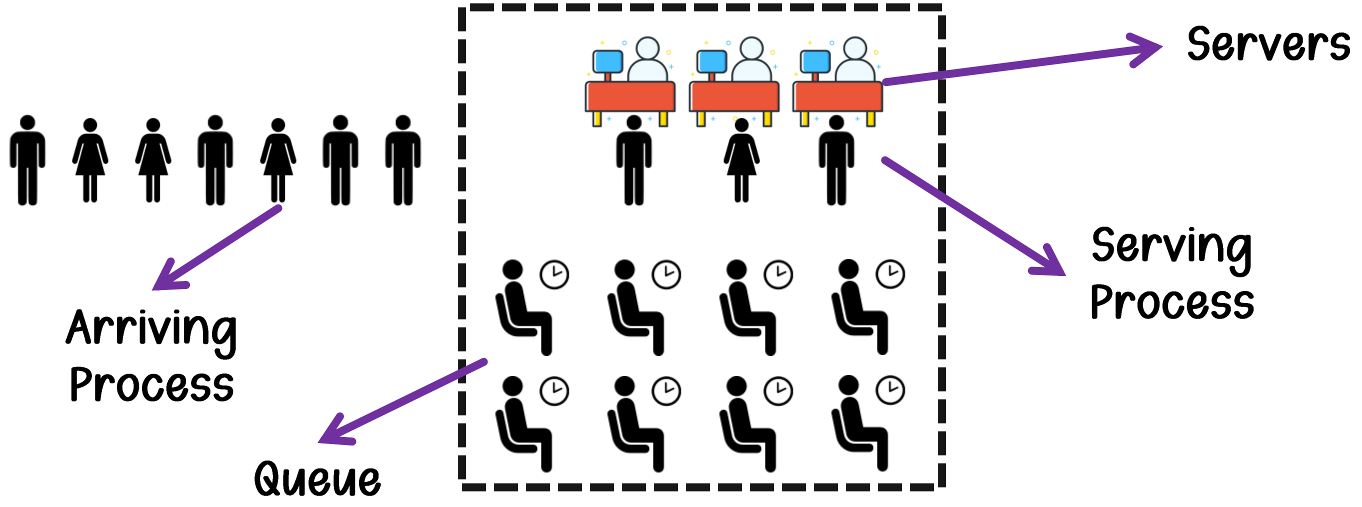
Queuing Theory

Section: [2.3]

Birth–Death Processes

Queuing System is Stochastic System

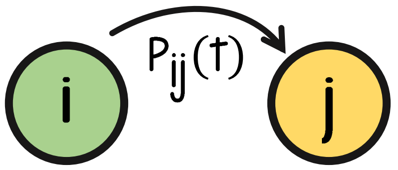
Queuing System



States of the System

The state of the queuing system at any time t is defined by the number of people in the system at that time t.

State Space $S = \{0, 1, 2, \cdots\}$



$P_{ij}(t)$ is the probability that j people will be present in the queuing system at time t, given that at time 0, i people are present.

Birth-Death Process

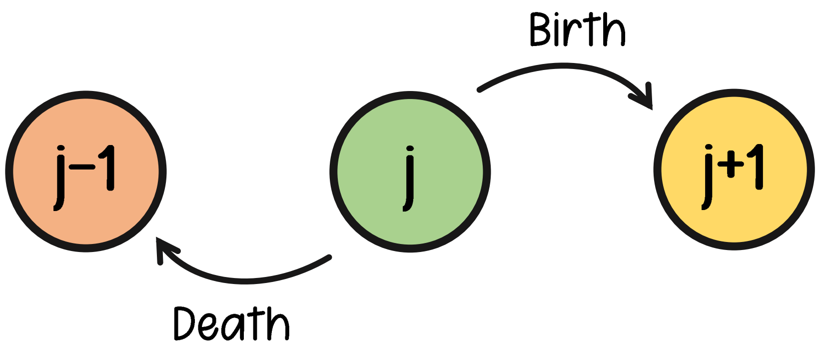
Definition

A **birth-death process** is a continuous-time stochastic process for which the system's state at any time is a nonnegative integer.

Three Laws

A **birth** (arrival) increases the system state by 1.

A **death** (service completion) decreases the system state by 1.



Births and deaths are **independent** of each other.

Steady State Probability

The **steady state (equilibrium) probability** of state j is denoted by π_j and equals the probability that there are j people in the system in the long run.

Rate Diagram — General Queuing System

Shape & Properties

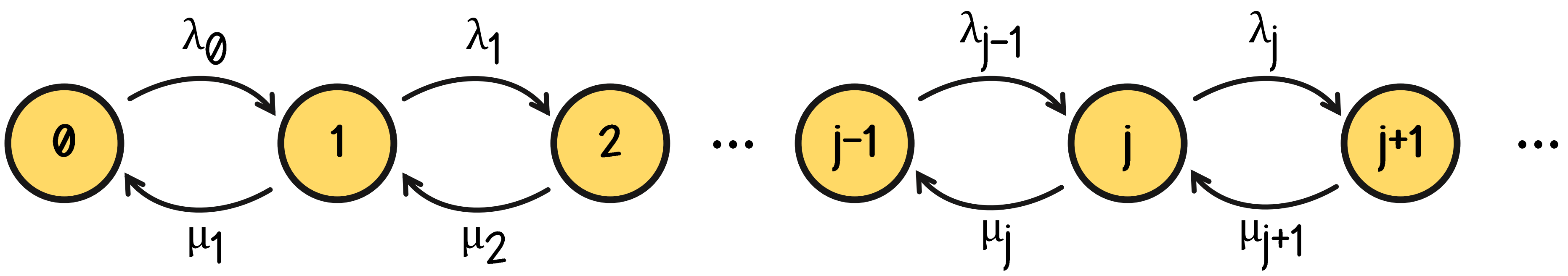
Each state is represented by a **circle**.

The number inside each circle indicates the number of customers in the system.

The transitions between states are represented by **arrows**.

The numbers put on the arrows indicate the transition rates between the states, either the arrival rates (λ_j) ore departure rate (μ_j).

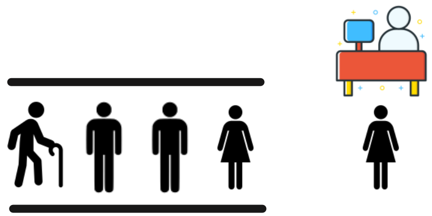
$\mu_0 = 0$



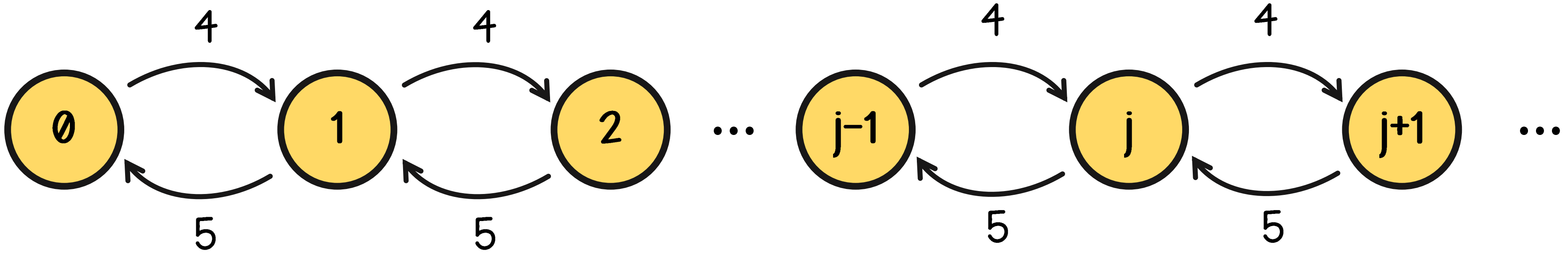
Rate Diagram — General Queuing System

Example Consider M/M/1/FIFO/∞/∞ queuing system in which interarrival times are exponential with $\lambda = 4$ and service times are exponential with $\mu = 5$.

$$\lambda_j = 4 \qquad \text{for } j = 0,1,2,\cdots$$

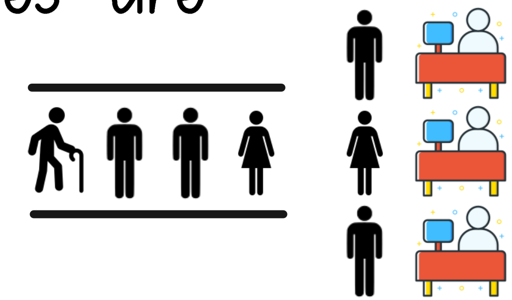


$$\mu_j = 5 \qquad \text{for } j = 1,2,3,\cdots \qquad \mu_0 = 0$$



Rate Diagram — General Queuing System

Example Consider M/M/3/FIFO/∞/∞ queuing system in which interarrival times are exponential with $\lambda = 4$ and service times are exponential with $\mu = 5$.



$\lambda_j = 4$

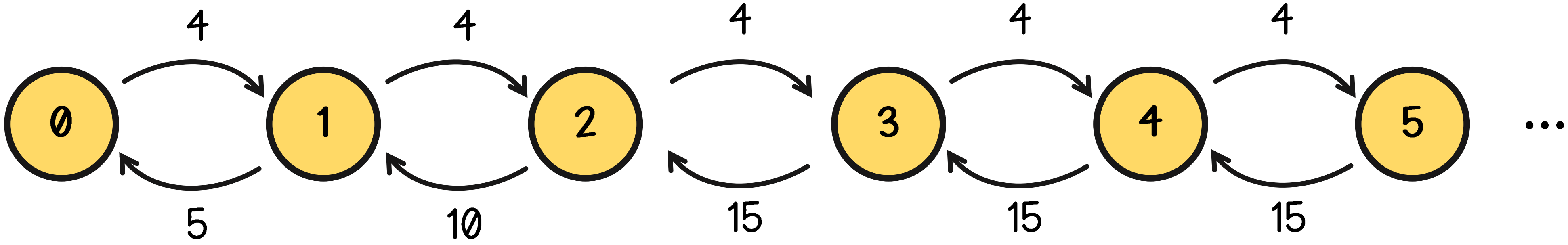
for $j = 0,1,2,\cdots$

$\mu_0 = 0$

$\mu_1 = 5$

$\mu_2 = 10$

$\mu_j = 15$ for $j = 3,4,5,\cdots$



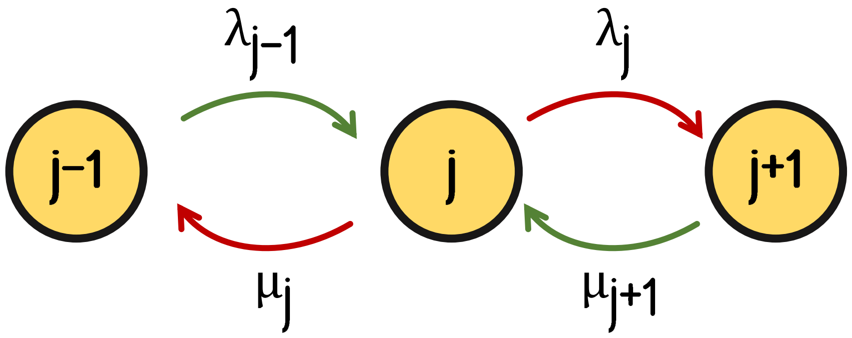
Computing Steady-State Probabilities

Balance Equation

The expected number of departures from state j in unit of time = The expected number of entrances to state j in unit of time

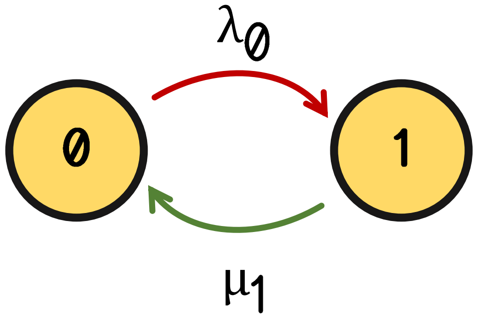
State $j \geq 1$

$$\pi_j \lambda_j + \pi_j \mu_j = \pi_{j-1} \lambda_{j-1} + \pi_{j+1} \mu_{j+1}$$



State $j = 0$

$$\pi_0 \lambda_0 = \pi_1 \mu_1$$



Computing Steady-State Probabilities

Formula For π_j

State 0 $\quad \pi_0\lambda_0 = \pi_1\mu_1 \qquad \pi_1 = \frac{\lambda_0}{\mu_1}\pi_0$

State 1 $\quad \pi_1\lambda_1 + \pi_1\mu_1 = \pi_0\lambda_0 + \pi_2\mu_2 \qquad \pi_2 = \frac{\lambda_0\lambda_1}{\mu_1\mu_2}\pi_0$

State 2 $\quad \pi_2\lambda_2 + \pi_2\mu_2 = \pi_1\lambda_1 + \pi_3\mu_3 \qquad \pi_3 = \frac{\lambda_0\lambda_1\lambda_2}{\mu_1\mu_2\mu_3}\pi_0$

$\vdots$

$$\pi_0\lambda_0 = \pi_1\mu_1$$

$$\pi_j\lambda_j + \pi_j\mu_j = \pi_{j-1}\lambda_{j-1} + \pi_{j+1}\mu_{j+1}$$

$$\pi_j = \frac{\lambda_0\lambda_1\lambda_2\cdots\lambda_{j-1}}{\mu_1\mu_2\mu_3\cdots\mu_j}\pi_0$$

Computing Steady-State Probabilities

What is π_0?

$$\pi_j = \underbrace{\boxed{\frac{\lambda_0\lambda_1\lambda_2\cdots\lambda_{j-1}}{\mu_1\mu_2\mu_3\cdots\mu_j}}}_{c_j}\pi_0$$

$$c_1 = \frac{\lambda_0}{\mu_1}$$

$$c_2 = \frac{\lambda_0\pi_1}{\mu_1\mu_2} = \frac{\lambda_1}{\mu_2}c_1$$

$$c_3 = \frac{\lambda_0\lambda_1\lambda_2}{\mu_1\mu_2\mu_3} = \frac{\lambda_2}{\mu_3}c_2$$

$$\vdots$$

$$c_j = \frac{\lambda_{j-1}}{\mu_j}c_{j-1}$$

$\therefore\ \pi_j = c_j\,\pi_0$ for $j \geq 1$

$$\pi_0 + \pi_1 + \pi_2 + \cdots = 1$$

for any stochastic system

$$\pi_0 + c_1\pi_0 + c_2\pi_0 + \cdots = 1$$

$$\pi_0(1 + c_1 + c_2 + \cdots) = 1$$

$$\pi_0 = \frac{1}{1 + \boxed{\sum_{j=1}^{\infty} c_j}}$$

Convergent?

Steady-state distribution exists.

Divergent?

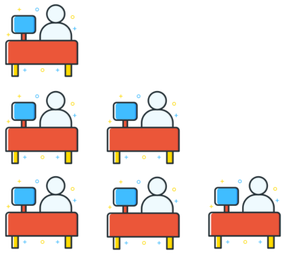
Steady-state distribution does not exist.

Computing Steady-State Probabilities

Example A grocery operates with three check-out counters. The manager uses the following schedule to determine the number of counters in operation, depending on the number of customers in store:

Customers arrive in the counters area according to a Poisson distribution with mean rate 10 customers per hour. The average check-out time per customer is exponential with mean 12 minutes. Determine the steady-state probability π_n of n customers in check-out area.

# customers in store	# counters in operation
1 2 3	
4 5 6	
7 …	



$\lambda_j = 10$ for all $j = 0,1,2,\cdots$

$\mu_0 = 0$

$\mu_1 = \mu_2 = \mu_3 = 5$

$\mu_4 = \mu_5 = \mu_6 = 10$

$\mu_j = 15$ for all $j = 7,8,9,\cdots$

$$\frac{1}{\mu} = 12 \text{ minutes/customer} \qquad \mu = \frac{60}{12} = 5 \text{ customers/hour}$$

Computing Steady-State Probabilities

Example $\pi_j = c_j \pi_0$ $\quad c_j = \frac{\lambda_{j-1}}{\mu_j} c_{j-1}$ $\quad \pi_0 = \frac{1}{1 + \sum_{j=1}^{\infty} c_j}$

$\lambda_j = 10$ for all $j = 0,1,2,\cdots$

$\mu_0 = 0$

$\mu_1 = \mu_2 = \mu_3 = 5$

$\mu_4 = \mu_5 = \mu_6 = 10$

$\mu_j = 15$ for all $j = 7,8,9,\cdots$

$$c_1 = \frac{\lambda_0}{\mu_1} = \frac{10}{5} = 2$$

$$c_2 = \frac{\lambda_1}{\mu_2} c_1 = \frac{10}{5} \cdot 2 = 4$$

$$c_3 = \frac{\lambda_2}{\mu_3} c_2 = \frac{10}{5} \cdot 4 = 8$$

$$c_4 = \frac{\lambda_3}{\mu_4} c_3 = \frac{10}{10} \cdot 8 = 8$$

$$c_5 = \frac{\lambda_4}{\mu_5} c_4 = \frac{10}{10} \cdot 8 = 8$$

$$c_6 = \frac{\lambda_5}{\mu_6} c_5 = \frac{10}{10} \cdot 8 = 8$$

$$c_7 = \frac{\lambda_6}{\mu_7} c_6 = \frac{10}{15} \cdot 8 = \left(\frac{2}{3}\right) \cdot 8$$

$$c_8 = \frac{\lambda_7}{\mu_8} c_7 = \frac{10}{15} \cdot \left(\frac{2}{3}\right) \cdot 8 = \left(\frac{2}{3}\right)^2 \cdot 8$$

$$c_j = 8 \cdot \left(\frac{2}{3}\right)^{j-6} \quad \text{for } j = 7,8,9,\cdots$$

Computing Steady-State Probabilities

Example $\pi_j = c_j \pi_0$ $\quad \pi_0 = \dfrac{1}{1 + \sum_{j=1}^{\infty} c_j}$

$c_1 = 2 \quad c_2 = 4$

$c_3 = c_4 = c_5 = c_6 = 8$

$c_j = 8 \cdot \left(\frac{2}{3}\right)^{j-6}$ for $j = 7, 8, 9, \cdots$

$$\sum_{j=1}^{\infty} c_j = c_1 + c_2 + c_3 + c_4 + c_5 + c_6 + c_7 + \cdots$$

$$= 2 + 4 + 8 + 8 + 8 + 8 + \boxed{\sum_{j=7}^{\infty} 8 \cdot \left(\frac{2}{3}\right)^{j-6}}$$

Geometric series $\quad r = \frac{2}{3} \quad |r| < 1$ (Convergent)

$$\sum_{j=7}^{\infty} 8 \cdot \left(\frac{2}{3}\right)^{j-6} = \sum_{j=0}^{\infty} 8 \cdot \left(\frac{2}{3}\right)^{j+1} = \frac{1^{st} \text{ term}}{1 - r}$$

$$= \frac{8 \cdot \frac{2}{3}}{1 - \frac{2}{3}} = 16$$

$$= 2 + 4 + 8 + 8 + 8 + 8 + 16$$

$$= 54$$

$$\therefore \pi_0 = \frac{1}{1 + 54} = \frac{1}{55}$$

Computing Steady-State Probabilities

Example

$$\pi_j = c_j \, \pi_0 \qquad \pi_0 = \frac{1}{55}$$

$c_1=2 \qquad c_2=4$

$c_3=c_4=c_5=c_6=8$

$c_j=8\cdot\left(\frac{2}{3}\right)^{j-6}$ for $j=7,8,9,\cdots$

$$\pi_1 = 2 \times \frac{1}{55} = \frac{2}{55}$$

$$\pi_2 = 4 \times \frac{1}{55} = \frac{4}{55}$$

$$\pi_3 = 8 \times \frac{1}{55} = \frac{8}{55} = \pi_4 = \pi_5 = \pi_6$$

$$\pi_j = 8\cdot\left(\frac{2}{3}\right)^{j-6} \times \frac{1}{55} = \frac{8}{55}\cdot\left(\frac{2}{3}\right)^{j-6} \quad \text{for } j=7,8,9,\cdots$$

Course: Applied Probability

Chapter: [2]

Queuing Theory

Section: [2.4]

The M/M/1/GD/∞/∞ Queuing System and the Queuing Formula $L = \lambda W$

Some Interesting Values in Queuing Systems

π_j steady state probability that there are j customers in the system.

L average number of customers present in the system.

L_q average number of customers waiting in line (queue).

L_s average number of customers in service.

W average time a customer spends in the system.

W_q average time a customer spends in line (queue).

W_s average time a customer spends in service.

Steady-State Probabilities for M/M/1 Model

Derivation

An M/M/1 queuing system can be modeled as a birth-death process with the following parameters.

$$\lambda_j = \lambda \qquad \text{for } j=0,1,2,\cdots$$

$$\mu_0 = 0$$

$$\mu_j = \mu \qquad \text{for } j=1,2,3,\cdots$$

c_j and π_j

$$c_j = \frac{\lambda_0}{\mu_1}\frac{\lambda_1}{\mu_2}\frac{\lambda_2}{\mu_3}\cdots\frac{\lambda_{j-1}}{\mu_j} = \left(\frac{\lambda}{\mu}\right)^j = \rho^j \quad \text{where} \quad \rho = \frac{\lambda}{\mu}$$

traffic intensity

$$\pi_j = c_j\pi_0 = \rho^j\pi_0$$

Steady-State Probabilities for M/M/1 Model

Steady State Probabilities

$$\pi_j = \rho^j \pi_0 \quad \text{where} \quad \rho = \frac{\lambda}{\mu}$$

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \cdots = 1$$

$$\pi_0 + \rho\pi_0 + \rho^2\pi_0 + \rho^3\pi_0 + \cdots = 1$$

$$\pi_0 \left(1 + \rho + \rho^2 + \rho^3 + \cdots\right) = 1$$

geometric series with $r = \rho$

converges to $\frac{1}{1-\rho}$ if $\rho < 1$

$$\pi_0 \left(\frac{1}{1-\rho}\right) = 1 \qquad \pi_0 = 1 - \rho$$

$$\therefore \pi_j = \rho^j(1-\rho) \quad \text{if } 0 \le \rho < 1$$

$$\boxed{0 \le \lambda < \mu}$$

Throughout the rest of this section, we assume that $\rho < 1$

Average Number of Customers in the System

Derivation of L

$$L = 0\cdot\pi_0 + 1\cdot\pi_1 + 2\cdot\pi_2 + 3\cdot\pi_3 + \cdots = \sum_{j=0}^{\infty} j\cdot\pi_j = \sum_{j=0}^{\infty} j\cdot\rho^j(1-\rho)$$

$$= (1-\rho)\boxed{\sum_{j=0}^{\infty} j\cdot\rho^j}$$

$$= (1-\rho)\frac{\rho}{(1-\rho)^2}$$

$$= \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

$$S = \rho + 2\rho^2 + 3\rho^3 + \cdots$$

$$\rho S = \rho^2 + 2\rho^3 + 3\rho^4 + \cdots \quad -$$

$$S - \rho S = \rho + \rho^2 + \rho^3 + \cdots$$

$$S(1-\rho) = \frac{\rho}{1-\rho}$$

$$\boxed{S = \frac{\rho}{(1-\rho)^2}}$$

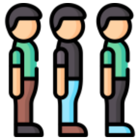
Average Number of Customers in Queue and Service

Derivation of L_q

$$L_q = 0 \cdot \pi_0 + 0 \cdot \pi_1 + 1 \cdot \pi_2 + 2 \cdot \pi_3 + \cdots$$

$$= \sum_{j=1}^{\infty} (j-1)\pi_j = \sum_{j=1}^{\infty} j\pi_j - \sum_{j=1}^{\infty} \pi_j$$

$$= L - (1 - \pi_0) = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

		
0	0	0
1	1	0
2	1	1
3	1	2
j	1	j-1

Derivation of L_s

$$L = L_q + L_s$$

$$L_s = L - L_q = \frac{\rho}{1-\rho} - \frac{\rho^2}{1-\rho} = \rho = \frac{\lambda}{\mu}$$

Little's Queuing Formula

Theorem For any queuing system in which a steady-state distribution exists, the following relations hold.

$$L = \lambda W \qquad L_q = \lambda W_q \qquad L_s = \lambda W_s \qquad W = W_q + W_s$$

M/M/1

$$W = \frac{1}{\mu - \lambda} \qquad W_q = \frac{\lambda}{\mu(\mu - \lambda)} \qquad W_s = \frac{1}{\mu}$$

Examples on M/M/1 Model

Example An average of 10 cars per hour arrive at a single-server drive-in teller. Assume that the average service time for each customer is 4 minutes, and both interarrival times and service times are exponential. Answer the following questions.

$$\lambda = 10$$

$$\mu = \frac{60}{4} = 15$$

$$\rho = \frac{\lambda}{\mu} = \frac{10}{15} = \frac{2}{3}$$

a) What is the probability that the teller is idle?

$$\pi_0 = 1 - \rho = 1 - \frac{2}{3} = \frac{1}{3}$$

b) What is the average number of cars waiting in line for the teller? (A car that is being served is not considered to be waiting in line.)

$$L_q = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{10^2}{15 \times 5} = \frac{4}{3} \text{ customers}$$

Examples on M/M/1 Model

Example An average of 10 cars per hour arrive at a single-server drive-in teller. Assume that the average service time for each customer is 4 minutes, and both interarrival times and service times are exponential. Answer the following questions.

$$\lambda = 10$$

$$\mu = \frac{60}{4} = 15$$

$$\rho = \frac{\lambda}{\mu} = \frac{10}{15} = \frac{2}{3}$$

c) What is the average amount of time a drive-in customer spends in the bank parking lot (including time in service)?

$$W = \frac{L}{\lambda} = \frac{1}{\mu - \lambda} = \frac{1}{5} \text{ hour}$$

$$= 12 \text{ minutes}$$

Examples on M/M/1 Model

Example For an M/M/1 queuing system, suppose that both λ and μ are doubled.

$$\lambda^* = 2\lambda$$

$$\mu^* = 2\mu$$

$$\rho^* = \frac{\lambda^*}{\mu^*} = \frac{2\lambda}{2\mu} = \rho$$

a) How is L changed?

$$L^* = \frac{\rho^*}{1-\rho^*} = \frac{\rho}{1-\rho} = L$$ unchanged

b) How is W changed?

$$W^* = \frac{L^*}{\lambda^*} = \frac{L}{2\lambda} = \frac{1}{2}W$$ cut in half

c) How is the steady-state probability distribution changed?

$$\pi_j^* = \rho^{*j}(1-\rho^*) = \rho^j(1-\rho) = \pi_j$$ unchanged

Course: Applied Probability

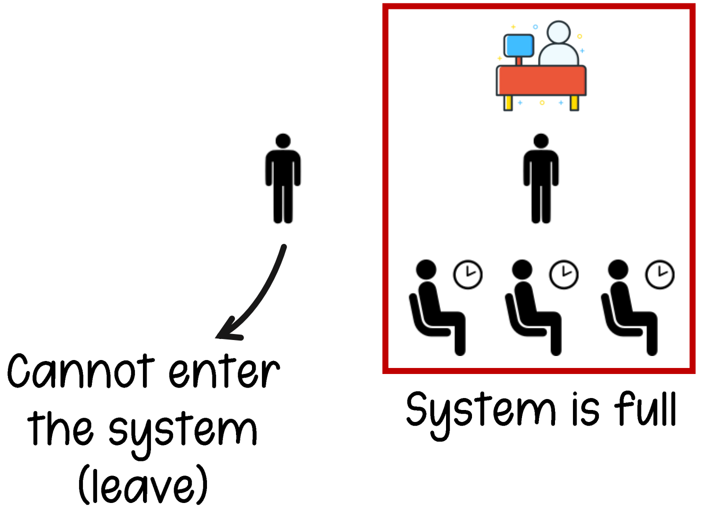
Chapter: [2]
Queuing Theory

Section: [2.5]
The M/M/1/GD/c/∞ Queuing System

Understanding the M/M/1/GD/c/∞ Model

Finite Capacity "c"

What does it mean that the capacity of a queuing system is c=4?



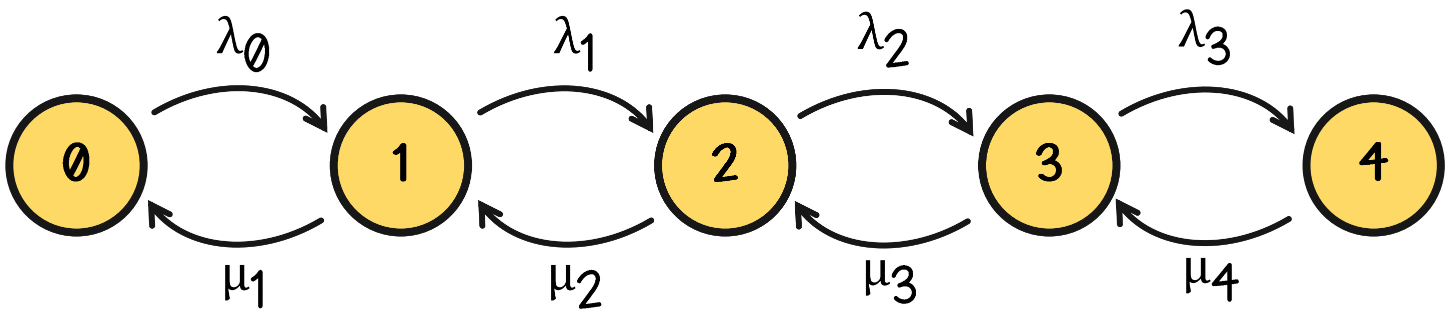
Parameters

$\lambda_j = \lambda$ for j=0,1,2,⋯,c−1

$\lambda_c = 0$

$\mu_0 = 0$

$\mu_j = \mu$ for j=1,2,3,⋯,c



Steady-State Probabilities for M/M/1/GD/c/∞ Model

c_j and π_j

$$c_j = \frac{\lambda_0}{\mu_1}\frac{\lambda_1}{\mu_2}\frac{\lambda_2}{\mu_3}\cdots\frac{\lambda_{j-1}}{\mu_j} = \left(\frac{\lambda}{\mu}\right)^n = \rho^n \quad \text{where} \quad \rho = \frac{\lambda}{\mu}$$

$$\pi_j = c_j\pi_0 = \rho^j\pi_0 \qquad \text{for } j=1,2,3,\cdots,c$$

Steady State Probabilities

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \cdots + \pi_c = 1$$

$$\pi_0 + \rho\pi_0 + \rho^2\pi_0 + \rho^3\pi_0 + \cdots + \rho^c\pi_0 = 1$$

$$\pi_0\left(1 + \rho + \rho^2 + \rho^3 + \cdots + \rho^c\right) = 1 \qquad \text{finite geometric series}$$

$$\pi_0\left(\frac{1-\rho^{c+1}}{1-\rho}\right) = 1 \qquad \pi_0 = \frac{1-\rho}{1-\rho^{c+1}} \quad \text{if } \rho \neq 1 \qquad \boxed{\lambda \neq \mu}$$

$$\therefore\ \pi_j = \frac{\rho^j(1-\rho)}{1-\rho^{c+1}} \quad \text{if } \rho \neq 1 \qquad \text{for } j=1,2,3,\cdots,c$$

Steady-State Probabilities for M/M/1/GD/c/∞ Model

Steady State Probabilities if $\rho = 1$

$$c_j = \rho^j = 1 \qquad \pi_j = c_j \pi_0 = \pi_0 \quad \text{for } j=1,2,3,\cdots,c$$

$$\underbrace{\pi_0 + \pi_1 + \pi_2 + \pi_3 + \cdots + \pi_c}_{(c+1)\text{ terms}} = 1$$

$$\pi_0 + \pi_0 + \pi_0 + \pi_0 + \cdots + \pi_0 = 1$$

$$\pi_0 (c+1) = 1 \qquad \pi_0 = \frac{1}{c+1} \quad \text{if } \rho = 1$$

$$\therefore \pi_j = \begin{cases} \dfrac{\rho^j (1-\rho)}{1-\rho^{c+1}} & : \quad \rho \neq 1 \ (\lambda \neq \mu) \\ \dfrac{1}{c+1} & : \quad \rho = 1 \ (\lambda = \mu) \end{cases} \qquad \text{for } j=1,2,3,\cdots,c$$

Rate of Entrance

Note In $M/M/1/GD/c/\infty$ not all arrivals will enter the system.

average number of customers who actually enter the system (Rate of Entrance)	=	average number of customers who arrive to the system (Arrival Rate)	−	average number of customers who cannot enter the system
λ^*		λ		$\lambda\pi_c$

$$\therefore\ \lambda^* = \lambda - \lambda\pi_c = \lambda(1 - \pi_c)$$

Average Number of Customers in the System

Derivation of L

If $\rho \neq 1$

$$L = 0\cdot\pi_0 + 1\cdot\pi_1 + 2\cdot\pi_2 + 3\cdot\pi_3 + \cdots + c\cdot\pi_c = \sum_{j=0}^{c} j\cdot\pi_j$$

$$= \sum_{j=0}^{c} \frac{j\cdot\rho^j(1-\rho)}{1-\rho^{c+1}} = \frac{\rho\left[1-(c+1)\rho^c + c\rho^{c+1}\right]}{\left(1-\rho^{c+1}\right)(1-\rho)}$$

If $\rho = 1$

$$L = 0\cdot\pi_0 + 1\cdot\pi_0 + 2\cdot\pi_0 + 3\cdot\pi_0 + \cdots + c\cdot\pi_0 = \pi_0 \cdot \sum_{j=0}^{c} j$$

$$= \frac{1}{1+c}\cdot\frac{c(c+1)}{2} = \frac{c}{2}$$

Average Number of Customers in Service and Queue

Derivation of L_s

$$L_s = 0 \cdot \pi_0 + 1 \cdot (\pi_1 + \pi_2 + \pi_3 + \cdots + \pi_c)$$
$$= 1 - \pi_0$$

Calculating L_q

$$L = L_q + L_s \qquad L_q = L - L_s$$

Average Waiting Time

Using Little's Queuing formula $L=\lambda W$:

$$W = \frac{L}{\lambda^*} = \frac{L}{\lambda(1-\pi_c)} \qquad W_q = \frac{L_q}{\lambda^*} = \frac{L_q}{\lambda(1-\pi_c)} \qquad W_s = \frac{L_s}{\lambda^*} = \frac{L_s}{\lambda(1-\pi_c)}$$

0	0
1	1
2	1
3	1
⋮	⋮
c	1

Example on the M/M/1/GD/c/∞ Model

Barber Shop

1 server capacity

A one-man barber shop has a total of 10 seats. Interarrival times are exponentially distributed, and an average of 20 prospective customers arrive each hour at the shop. Those customers who find the shop full do not enter. The barber takes an average of 12 minutes to cut each customer's hair. Haircut times are exponentially distributed.

a) What is the probability that the barber is busy?

$$c = 10$$

$$\lambda = 20$$

$$\mu = \frac{60}{12} = 5$$

$$\rho = \frac{\lambda}{\mu} = 4$$

Probability Of Busy

$$= \pi_1 + \pi_2 + \cdots + \pi_{10} = 1 - \pi_0$$

$$= 1 - \frac{1-\rho}{1-\rho^{c+1}} = 1 - \frac{1-4}{1-4^{11}} \approx 0.999$$

Example on the M/M/1/GD/c/∞ Model

Barber Shop

b) On the average, how much time will be spent in the shop by a customer who enters?

$$c = 10 \qquad \lambda = 20 \qquad \mu = \frac{60}{12} = 5 \qquad \rho = \frac{\lambda}{\mu} = 4$$

$$L = \frac{\rho\left[1-(c+1)\rho^{c}+c\rho^{c+1}\right]}{(1-\rho^{c+1})(1-\rho)} = \frac{4\left[1-(11)4^{10}+10\cdot 4^{11}\right]}{(1-4^{11})(1-4)} \approx 9.667$$

$$\pi_j = \frac{\rho^j(1-\rho)}{1-\rho^{c+1}} \qquad \pi_{10} = \frac{4^{10}(1-4)}{1-4^{11}} \approx 0.75$$

$$W = \frac{L}{\lambda^*} = \frac{L}{\lambda(1-\pi_c)} = \frac{9.667}{20(1-0.75)} \approx 1.933 \text{ hours}$$

Course: Applied Probability

Chapter: [2]
Queuing Theory

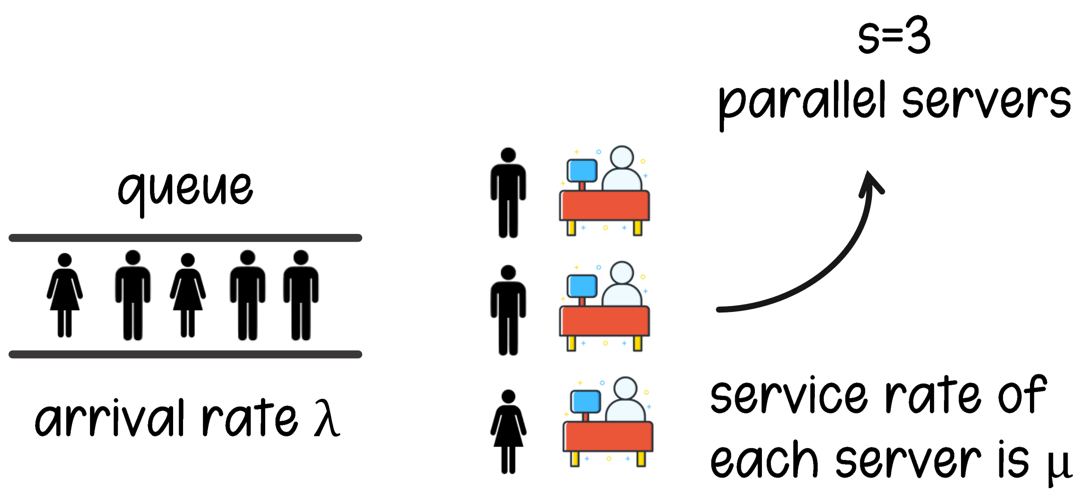
Section: [2.6]
The M/M/s/GD/∞/∞ Queuing System

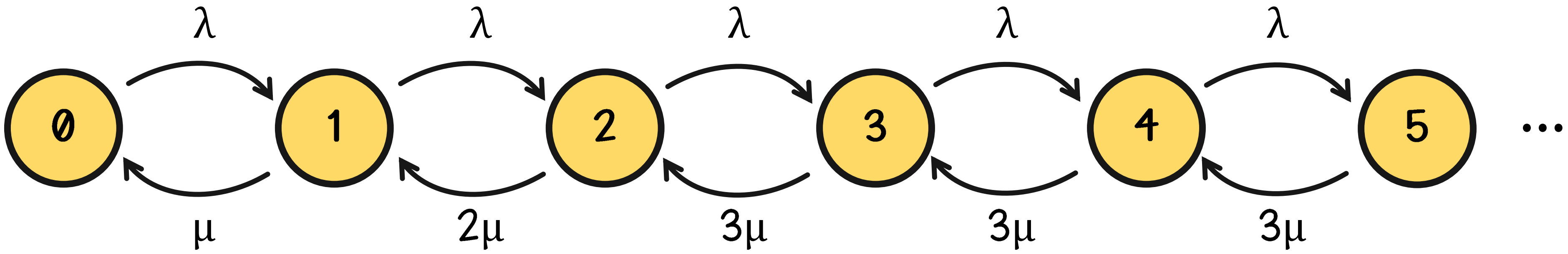
Understanding the M/M/s/GD/ ∞ /∞ Model

Parameters of Parallel Servers

$$\lambda_j = \lambda \qquad \text{for } j=0,1,2,\cdots,c-1$$

$$\mu_j=\begin{cases} j\mu & : \quad j=0,1,2,\cdots,s \\ s\mu & : \quad j=s+1,s+2,\cdots \end{cases}$$





Steady-State Probabilities for M/M/s Model

Formulas for π_j

In M/M/s model, we define $\rho = \frac{\lambda}{s\,\mu}$.

Steady-state probabilities in this model are exist if $\rho < 1$.

$$\pi_0 = \frac{1}{\frac{(s\rho)^s}{s!(1-\rho)} + \sum_{i=0}^{s-1} \frac{(s\rho)^i}{i!}}$$

$$\pi_j = \begin{cases} \frac{(s\rho)^j \pi_0}{j!} & : \quad j=1,2,\cdots,s \\ \frac{(s\rho)^j \pi_0}{s! \cdot s^{j-s}} & : \quad j=s+1,s+2,\cdots \end{cases}$$

$$P\begin{pmatrix}\text{all servers} \\ \text{are busy}\end{pmatrix} = P(j \geq s)$$

$$= \frac{(s\rho)^s \pi_0}{s!(1-\rho)}$$

$$= 1 - P\begin{pmatrix}\text{at least} \\ \text{one server} \\ \text{is idle}\end{pmatrix}$$

$$\pi_0 = \frac{P(j \geq s)\, s!(1-\rho)}{(s\rho)^s}$$

$P(j \geq s)$ for M/M/s Model

ρ	$s = 2$	$s = 3$	$s = 4$	$s = 5$	$s = 6$	$s = 7$
.10	.02	.00	.00	.00	.00	.00
.20	.07	.02	.00	.00	.00	.00
.30	.14	.07	.04	.02	.01	.00
.40	.23	.14	.09	.06	.04	.03
.50	.33	.24	.17	.13	.10	.08
.55	.39	.29	.23	.18	.14	.11
.60	.45	.35	.29	.24	.20	.17
.65	.51	.42	.35	.30	.26	.21
.70	.57	.51	.43	.38	.34	.30
.75	.64	.57	.51	.46	.42	.39
.80	.71	.65	.60	.55	.52	.49
.85	.78	.73	.69	.65	.62	.60
.90	.85	.83	.79	.76	.74	.72
.95	.92	.91	.89	.88	.87	.85

Queuing Formulas for M/M/s Model

Average Number of Customers

$$L_q = \frac{P(j \geq s)\rho}{1-\rho} \qquad L_s = \frac{\lambda}{\mu} = s\rho \qquad L = L_s + L_q$$

Average Waiting Times

$$W_q = \frac{L_q}{\lambda} = \frac{P(j \geq s)}{s\mu - \lambda} \qquad W_s = \frac{L_s}{\lambda} = \frac{1}{\mu} \qquad W = W_s + W_q$$

Example on M/M/s Model

Bank with Two Tellers

two servers arrival rate

Consider a bank with two tellers. An average of 80 customers per hour arrive at the bank and wait in a single line for an idle teller. The average time it takes to serve a customer is 1.2 minutes. Assume that interarrival times and service times are exponential. Determine

service rate

a) The expected number of customers present in the bank.

$$L_s = \frac{\lambda}{\mu} = \frac{80}{50} = 1.6 \quad \text{customers}$$

$$L_q = \frac{P(j \geq 2)\rho}{1-\rho} = \frac{(0.71)(0.8)}{0.2} = 2.84 \quad \text{customers}$$

$$L = L_s + L_q = 1.6 + 2.84 = 4.44 \quad \text{customers}$$

$s=2$

$\lambda=80$

$$\mu = \frac{60}{1.2} = 50$$

$$\rho = \frac{\lambda}{s\,\mu} = 0.8$$

$P(j \geq 2) = 0.71$

Example on M/M/s Model

Bank with Two Tellers

b) The expected length of time a customer spends in the bank.

$$W=\frac{L}{\lambda}=\frac{4.44}{80}=0.055 \text{ hours} = 3.3 \text{ minutes}$$

c) What fraction of time both servers are idle.

$$\pi_0=\frac{1}{\frac{(s\rho)^s}{s!(1-\rho)}+\sum_{i=0}^{s-1}\frac{(s\rho)^i}{i!}}=\frac{1}{\frac{(2\times 0.8)^2}{2!\times 0.2}+\left(1+\frac{2\times 0.8}{1}\right)}\approx 0.111$$

$$\pi_0=\frac{P(j\geq s)s!(1-\rho)}{(s\rho)^s}=\frac{P(j\geq 2)\times 2!\times 0.2}{(2\times 0.8)^2}\approx 0.111$$

$s=2$

$\lambda=80$

$$\mu=\frac{60}{1.2}=50$$

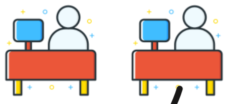
$$\rho=\frac{\lambda}{s\,\mu}=0.8$$

$P(j\geq 2)=0.71$

Example on M/M/s Model

Bank with Two Tellers

d) The fraction of time that a particular teller is idle.



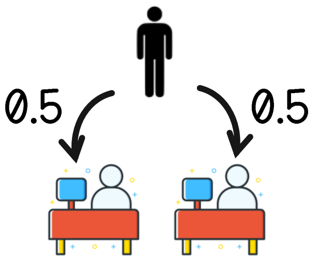
idle if both tellers are idle

$\downarrow$

no customers in the system

$\downarrow$

π_0



busy idle

$\downarrow$

1 customer in the system

$\downarrow$

$0.5\pi_1$

$$\pi_j = \frac{(s\rho)^j \pi_0}{j!} \qquad \pi_1 = \frac{(2\times 0.8)^1 \times 0.111}{1!} \approx 0.178$$

$$P\left(\begin{matrix}\text{particular teller}\\ \text{is idle}\end{matrix}\right) = \pi_0 + 0.5\pi_1 = 0.20$$

$s=2$

$\lambda=80$

$\mu = \frac{60}{1.2} = 50$

$\rho = \frac{\lambda}{s\,\mu} = 0.8$

$P(j\geq 2) = 0.71$

$\pi_0 = 0.111$