22

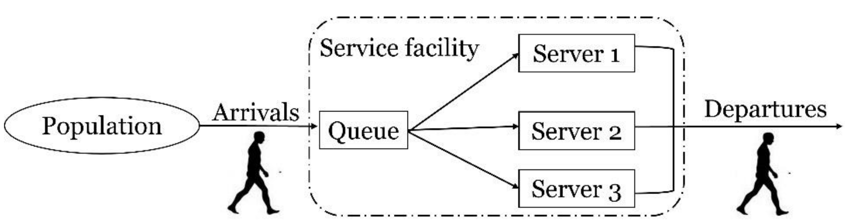
QUEUEING THEORY & MODELS

CONTENTS

Queueing theory is a mathematical study of waiting lines/queues, used to predict queue lengths and waiting times. It's a branch of operations research, applied in various fields like telecommunication, traffic engineering, computing, project management and industrial engineering (designing factories, shops, offices, hospitals).

Queueing theory has its roots in Agner Krarup Erlang's research on telephone exchange systems and has since been widely applied to optimize resource allocation and improve efficiency in various industries.

Example: Consider a bank where customers wait in line to be served by tellers. Understanding the Queueing behavior can help the bank manage wait times and improve customer satisfaction.



The *work conservation rule* applies: servers cannot be idle when customers are waiting.

ELEMENTS OF A QUEUEING SYSTEM

A typical Queueing system consists of several key components:

- *Arrival Process:* The way customers arrive at the queue, usually modeled by the Poisson distribution.
- *Service Process:* How customers are served, often modeled by the exponential distribution.
- *Queue Discipline:* The order in which customers are served, such as First-In-First-Out (FIFO) or Last-In-First-Out (LIFO).
- *Service Channels:* The number of servers available to serve customers.
- *Customer Behavior:* Actions like balking (refusing to join a long queue), reneging (leaving a queue after waiting too long), or jockeying (customers changing queue to reduce the waiting time).

Types of Service Order/Queueing Disciplines

This is how an item or customer is served on arrival at a service facility:

- *FIFO (First In, First Out):* Customers are served in the order they arrive.

- *LIFO (Last In, First Out):* Last customer to arrive is served first.
- *SIRO (Service In Random Order):* Customers are served in random order.
- *PRI (Priority Service):* Customers are prioritized based on their importance or urgency.

Important Distributions in Queueing Theory

These distributions are fundamental to queueing theory, as they help model and analyze the behavior of queues.

i. *Poisson Distribution:* Models the number of arrivals in a fixed time period. Assumes arrivals are independent and uniform.
ii. *Exponential Distribution:* Models the time between arrivals or service times. Assumes exponential decay in the rate of arrivals or service.
iii. *Erlang-K Distribution:* Generalization of the exponential distribution. Models the sum of K independent exponential variables. Used for modeling service times with K servers in parallel.
iv. *Hyper-Exponential-K Distribution:* Generalization of the exponential distribution. Models the minimum of K independent exponential variables. Used for modeling service times with K servers in series.

Types of Queueing Models

There are several types of Queueing models, each suited to different scenarios:

a. Single Queue, Single Server (M/M/1)

This model features a single line and one server. It's one of the simplest and most common Queueing models.

Example: A customer service desk with one representative. Customers arrive at random (Poisson process) and are served one at a time (exponential service time).

b. Single Queue, Multiple Servers (M/M/c)

This model has a single line but multiple servers. It's used when a system can handle several customers simultaneously.

Example: A bank with multiple tellers serving customers from a single line.

c. Multiple Queues, Multiple Servers (M/M/c)

In this model, multiple lines lead to multiple servers.

Example: Supermarkets where each checkout counter has its own line.

Kendall's Notation

Kendall's notation is a standardized way to describe queueing models using the following notations:

$$(a/b)/c{:}\,(d/e)/f$$

a: Arrival pattern (e.g., M for Poisson).
b: Service pattern (e.g., M for exponential).
c: Number of servers.
d: System discipline (e.g., FIFO).
e: Maximum number of customers allowed in the system.
f: Size of the calling population.

Example: For a single server system with Poisson arrivals and exponential service times:

$$M/M/1$$

☑ Example 22.1

A petrol station has one pump. Cars arrive according to a Poisson process and are served in order of arrival. The service time is exponentially distributed. Describe this system in Kendall's notation.

SOLUTION **tips**

Kendall’s Notation:

a = Arrival pattern = Poisson = M

Purchase the full book at:
https://unimath.5profz.com/

We donate 0.5% of the book sales every year to charity, forever. When you buy **University Mathematics (I & II)** *you are helping orphans and the less privileged.*