

**Mathematical Modeling:
Discrete Models and Markov Chains**

Wai-Ki CHING
Department of Mathematics
The University of Hong Kong

September 2025 (Beijing)

1 Mathematical Modeling: Discrete Models and Markov Chains

The Outline

This part aims at providing the basic knowledge in mathematical modeling, probability theory, Poisson process and Markov chain theory with some applications.

Some applications such as inventory systems and PageRank algorithm are discussed.

- (1) **Mathematical Modeling: Discrete Models and Difference Equations with Examples**
- (2) **Probability Theory, Poisson Process and Inventory Systems**

Mathematics is about proof.

Proof

證明

言 登 日 月

Words

Go up to

Sun

&

Mon

Figure 1.1: What is Mathematics?

- Mathematics is about “proving” a statement with some axioms/assumption and definitions.

Example:

Definition: On a **vector space** V , a **norm** is a function $\|\cdot\|$ from V to the set of **non-negative real numbers** such that

- (i) $\|\mathbf{x}\| > 0 \quad \forall \mathbf{x} \in V$ and $\mathbf{x} \neq \mathbf{0}$;
- (ii) $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\| \quad \forall \mathbf{x} \in V, \lambda \in \mathbb{R}$;
- (iii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in V$.

Theorem: ℓ_2 -norm: $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$ where $\mathbf{x} = [x_1, \dots, x_n]^T$, is a vector norm.

Proof: (i) We note that if $\mathbf{x} \neq \mathbf{0}$, then there is at least one $x_i \neq 0$. Thus

$$\left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}} \neq 0.$$

(ii) We note that

$$\|\lambda\mathbf{x}\|_2 = \left(\sum_{i=1}^n (\lambda x_i)^2\right)^{\frac{1}{2}} = |\lambda| \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}.$$

(iii) Moreover, we also have

$$\begin{aligned} \left(\left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^n y_i^2 \right)^{\frac{1}{2}} \right)^2 &= \left(\sum_{i=1}^n x_i^2 \right) + \left(\sum_{i=1}^n y_i^2 \right) + 2 \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n y_i^2 \right)^{\frac{1}{2}} \\ &\geq \left(\sum_{i=1}^n x_i^2 \right) + \left(\sum_{i=1}^n y_i^2 \right) + 2 \left(\sum_{i=1}^n |x_i| |y_i| \right) \\ &= \left(\sum_{i=1}^n (|x_i| + |y_i|)^2 \right) \geq \left(\sum_{i=1}^n (|x_i + y_i|)^2 \right). \end{aligned}$$

Hence the result follows.

(i) - (iii) are the assumptions that we made to define a norm (Why?).

- Mathematics is also a **language**. Language is the **dress of thought** (Samuel Johnson). Moreover, the **limits of my language** are the **limits of my world** (Ludwig Wittgenstein).

- Mathematics can be used to describe our real world and help us to have a better understanding of our universe. And **Mathematical Modeling Technique** is a very good candidate in the set of **transferable skills**.

Transferable Skills

- **Skills** can be used in a variety of jobs or occupations.
- **Transferable skills** are the **inventory of assets** that help you transition into and excel in a **new role**. They ensure the **robustness and longevity** of your career. Moreover, they allow you to more easily and readily **explore lateral dimensions** in your career and acquire added skills and expertise.
- While **highly specialized skills** may be essential to building your own personal competitive advantage and ensuring success in a **particular role or organization**, it is the **transferable skills** that ensure you **DO NOT become professionally redundant** or obsolete over the long term.

[Taken from <http://www.bayt.com/en/career-article-2721/>]

Transferable Mathematical Knowledge

- **Mathematical Models:** **Deterministic Models, Stochastic Models,** Statistical Models, etc.
- **Stochastic Processes:** **Markov Chain Process, Poisson Process,** Birth-and-Death Process, Diffusion Process, Levy Process, etc.
- **Computational Methods:** Iterative Methods, Numerical Algorithms, **Computer Simulations,** etc.
- **Data Mining Techniques:** **Data Modeling,** Classifications methods, Clustering Methods, etc.

(1) Mathematical Modeling: Discrete Models and Difference Equations

- Examples in Economics, Epidemic, Transportation Networks, etc.

(2) Probability Theory, Poisson Process and Inventory Systems

-Exponential Distribution and Poisson Distribution

-Poisson Process

-Examples in Inventory Models

We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances.

Isaac Newton (Wikipedia)

2 Mathematical Modeling

- A mathematical model ¹ is a description of a **system** using **mathematical concepts** and **language**. The **process** of developing a mathematical model is termed **mathematical modeling**.
- Mathematical models are used in the natural **Sciences** (such as Physics, Chemistry, Biology, Earth Science) and **Engineering** and **Technology** (such as Computer Science and Artificial Intelligence), as well as **Social Sciences** (such as Economics, Psychology, Sociology, Political Science), and **Management**.
- Physicists, engineers, statisticians, operations research analysts, and economists use mathematical models extensively. A model helps to **explain** a system, to study the effects of different components, and to make **predictions about behaviour** and rational **decisions**.

¹Taken from Wikipedia, the free encyclopedia. (http://en.wikipedia.org/wiki/Mathematical_model)

- Mathematical models can take many forms, including but not limited to **Dynamical Systems**, **Statistical Models**, **Differential Equations**, or **Game Theoretic Models** etc.
- In general, mathematical models may include logical models. The quality of a scientific field depends on how well the mathematical models developed on the theoretical side **agree** with the results of repeatable experiments. Lack of agreement between **theoretical mathematical models** and **experimental measurements** often leads to important advances as better theories are developed.
- All models are **wrong**, but some are useful (**George E. P. Box**).
- Any intelligent fool can make things bigger, more complex, and more violent. It takes a touch of genius – and a lot of courage – to move in the opposite direction (**Albert Einstein**)
- **KISS Principle**: Keep It Simple and Smart.

2.1 The Cobweb Model

The Cobweb model is an economic theory for studying price fluctuation and market equilibrium. Let p_t be the price, q_t be quantity at time t , and the model reads:

$$\begin{cases} q_t = \alpha + \beta p_{t-1} \\ p_t = \gamma - \delta q_t \end{cases}$$

where where $\alpha, \beta > 0, \gamma, \delta > 0$ and $q_0 > 0$ are given and therefore $p_0 = \gamma - \delta q_0$.

Solving the above two equations, we obtain two first-order linear difference equations:

$$q_t = \alpha + \beta\gamma - \beta\delta q_{t-1}$$

and

$$p_t = \gamma - \alpha\delta - \beta\delta p_{t-1}.$$

It is straightforward to solve for p_t and q_t , but the long-run behavior of the solutions are of special interest.

Since $\beta > 0, \delta > 0$, their product $\beta\delta > 0$, the solutions are thus always oscillatory with respect to time t .

As $t \rightarrow \infty$, the equilibrium point is

$$(p^*, q^*) = \left(\frac{\gamma - \alpha\delta}{1 + \beta\delta}, \frac{\alpha + \beta\gamma}{1 + \beta\delta} \right)$$

which is obtained by solving

$$\begin{cases} q = \alpha + \beta\gamma - \beta\delta q \\ p = \gamma - \alpha\delta - \beta\delta p. \end{cases}$$

(1) If $0 < \beta\delta < 1$, then the solutions $\{p_t\}$ and $\{q_t\}$ will converge to (p^*, q^*) .

(2) If $\beta\delta = 1$, the solutions will oscillate with finite magnitude.

(3) If $\beta\delta > 1$, the solutions will oscillate with infinite magnitude.

2.2 The SIR Model

Modeling the spread of an epidemic was issued by Kermack and McKendrick (1927). In a closed population of size N , at time n , we let

x_n be the population of the susceptible;

y_n be the population of the infective;

z_n be the population of the removal.

- The total population is assumed to be homogeneously mixed and is constant at any time n , i.e.,

$$x_n + y_n + z_n = N.$$

- Let β be the infection rate and γ be the removal rate. Then, the system dynamics can be described as follows:

$$\begin{cases} x_{n+1} = x_n - \beta \cdot x_n \cdot y_n \\ y_{n+1} = y_n + \beta \cdot x_n \cdot y_n - \gamma \cdot y_n \\ z_{n+1} = z_n + \gamma \cdot y_n \end{cases}$$

Suppose $\beta = 0.02$ and $\gamma = 0.2$.

- The population size $N = 100$ and initial at $n = 0$ there is one infective. Using the difference equations, we know within one week, all people will be infected.

n	x_n	y_n	z_n
0	99.0	1.0	0.0
1	97.0	2.8	0.2
2	91.6	7.6	0.8
3	77.7	20.1	2.2
4	46.5	47.2	6.3
5	2.6	81.7	15.8
6	0.0	69.6	30.4

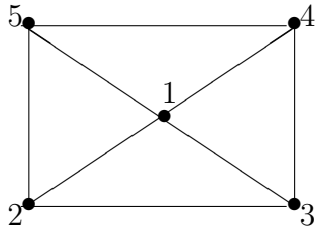


Figure 2.1: The random walker and the network.

2.3 A Random Walker (Traveler) Problem

Consider a random walker (traveler) in a network of cities as shown in Figure 2.1. The walker travels every day to the cities in the network.

- (i) Suppose that at each city i ($i = 1, 2, 3, 4, 5$), the probabilities of traveling to other adjacent cities are all equal.
- (ii) While the probability of staying at the same cite in the next move is assumed to be zero.
- (iii) At time 0 (day 0), the traveler is in **City 2**.
- **Question:** What is the probability that at **time 7 (day 7)** (one week later), the traveler is found in **City 1**?

●Observations:

We note that the probability of which city to visit tomorrow (time $t + 1$) will be fixed if the current city that the traveler visited (time t) is known.

Suppose today the traveler is in City 1 (time t), then with probability $1/4$ he will visit Cities 2,3,4 or 5 tomorrow (time $t + 1$).

If today the traveler is in City 2 (time t), then with probability $1/3$ he will visit Cities 1,3,or 5 tomorrow (time $t + 1$). But the probability of visiting City 4 is 0.

• Problem Formulation:

We define the **transition probability** Q_{ij} . Here Q_{ij} is the probability that the traveler will make a move to City i given the traveler is now in City j .

$$Q_{i1} = 1/4 \text{ for } i = 2, 3, 4, 5;$$

$$Q_{i2} = 1/3 \text{ for } i = 1, 3, 5; \quad Q_{i3} = 1/3 \text{ for } i = 1, 2, 4;$$

$$Q_{i4} = 1/3 \text{ for } i = 1, 3, 5; \quad Q_{i5} = 1/3 \text{ for } i = 1, 2, 4;$$

The remaining Q_{ij} are all zero. If we know the probabilities, then we write the matrix

$$Q = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \end{bmatrix}.$$

Let $P_i(t)$ be the probability that the traveler is in City i at time t . Here $i = 1, 2, 3, 4, 5$ and $t = 0, 1, 2, \dots$.

Now we have

$$\left\{ \begin{array}{l} P_1(t+1) = Q_{12}P_2(t) + Q_{13}P_3(t) + Q_{14}P_4(t) + Q_{15}P_5(t) \\ P_2(t+1) = Q_{21}P_1(t) + Q_{23}P_3(t) + Q_{25}P_5(t) \\ P_3(t+1) = Q_{31}P_1(t) + Q_{32}P_2(t) + Q_{34}P_4(t) \\ P_4(t+1) = Q_{41}P_1(t) + Q_{43}P_3(t) + Q_{45}P_5(t) \\ P_5(t+1) = Q_{51}P_1(t) + Q_{52}P_2(t) + Q_{54}P_4(t) \end{array} \right.$$

and

$$P_1(0) = P_3(0) = P_4(0) = P_5(0) = 0 \quad \text{and} \quad P_2(0) = 1.$$

Or

$$\left\{ \begin{array}{l} P_1(t+1) = \\ P_2(t+1) = 1/4P_1(t) \\ P_3(t+1) = 1/4P_1(t) +1/3P_2(t) \\ P_4(t+1) = 1/4P_1(t) +1/3P_3(t) \\ P_5(t+1) = 1/4P_1(t) +1/3P_2(t) \end{array} \right.$$

and

$$P_1(0) = P_3(0) = P_4(0) = P_5(0) = 0 \quad \text{and} \quad P_2(0) = 1.$$

- In matrix language, the relation can be written as follows:

$$\begin{bmatrix} P_1(t+1) \\ P_2(t+1) \\ P_3(t+1) \\ P_4(t+1) \\ P_5(t+1) \end{bmatrix} = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \\ 1/4 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/3 & 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ P_4(t) \\ P_5(t) \end{bmatrix}. \quad (2.1)$$

- If we let

$$\mathbf{P}(t) = [P_1(t) \ P_2(t) \ P_3(t) \ P_4(t) \ P_5(t)]^T$$

then we may write

$$\mathbf{P}(t+1) = Q\mathbf{P}(t)$$

and

$$\mathbf{P}(0) = [P_1(0) \ P_2(0) \ P_3(0) \ P_4(0) \ P_5(0)]^T = [0 \ 1 \ 0 \ 0 \ 0]^T$$

and we are asked to find $P_1(7)$ of $\mathbf{P}(7)$.

- Using MATLAB, we compute the followings:

P(1)	P(2)	P(3)	P(4)	P(5)	P(6)	P(7)
0.3333	0.2222	0.2593	0.2469	0.2510	0.2497	0.2501
0.0000	0.3056	0.1111	0.2376	0.1543	0.2095	0.1728
0.3333	0.0833	0.2593	0.1389	0.2202	0.1656	0.2021
0.0000	0.3056	0.1111	0.2377	0.1543	0.2095	0.1728
0.3333	0.0833	0.2593	0.1389	0.2202	0.1656	0.2021

The answer is then equal to 0.2501 roughly.

Remark 2.1: One may consider the long-run probabilities, i.e.,

$$\lim_{t \rightarrow \infty} P_i(t) = P_i, \quad i = 1, 2, 3, 4, 5.$$

Suppose they exist, then they must satisfy

$$\begin{cases} P_1 = & 1/3P_2 & +1/3P_3 & +1/3P_4 & +1/3P_5 \\ P_2 = 1/4P_1 & & +1/3P_3 & & +1/3P_5 \\ P_3 = 1/4P_1 & +1/3P_2 & & +1/3P_4 & \\ P_4 = 1/4P_1 & & +1/3P_3 & & +1/3P_5 \\ P_5 = 1/4P_1 & +1/3P_2 & & +1/3P_4 & \end{cases} \quad (2.2)$$

and

$$P_1 + P_2 + P_3 + P_4 + P_5 = 1.$$

- What is the meaning of P_i ? It means that when the traveling process has been continued for a long time (**in equilibrium**), the probability of finding the traveler in City i will be P_i

- Here by symmetry of the network, $P_2 = P_3 = P_4 = P_5 = P$, we have

$$P_1 + 4P = 1 \quad \text{and} \quad P_1 = 4/3P \quad (\text{the first equation}).$$

Hence, we have $P_1 = 1/4$ and $P_2 = P_3 = P_4 = P_5 = P = 3/16$.

$$\lim_{t \rightarrow \infty} \mathbf{P}(t) = \begin{bmatrix} 4 & 3 & 3 & 3 & 3 \\ 16 & 16 & 16 & 16 & 16 \end{bmatrix}^T.$$

- The traveler is “spending” 25% of his time in City 1. While the other four cities will share equally the remaining 75% of his visiting time.
- If we regard a city as a webpage, the traveler as an user in the internet and the edges as the links.

Then, Webpage 1 is the most important one (rank no. 1) because the user spends most of his time (25%) on this webpage.

The other four webpages have the same ranking. This is the key idea of Google’s PageRank algorithm.

Remark 2.2: In the example, we ignore the **transportation cost** which may affect the traveler's decision. Suppose the highways connecting the cities have different fees (costs) given in the following table.

City	1	2	3	4	5
1	∞	20	25	40	50
2	20	∞	20	∞	40
3	25	20	∞	25	∞
4	40	∞	25	∞	10
5	50	40	∞	10	∞

Question: How to model the probabilities in the matrix of Eq. (2.1)?

Let Q_{ij} be the probability that the traveler will move to City i tomorrow given that today he is in City j . Let C_{ij} be the cost of taking the highway from City i to City j (in the above table). What is an appropriate relation between Q_{ij} and C_{ij} ?

• One possible suggestion is to assume that Q_{ij} is a decreasing function in C_{ij} . For example,

(a) $Q_{ij} \propto \frac{1}{C_{ij}^2}$

(b) $Q_{ij} \propto e^{-C_{ij}}$

Suppose we adopt (a), then let us compute Q_{i1} for $i = 1, 2, 3, 4$. We have (for some $K > 0$),

$$Q_{21} = \frac{K}{20}, \quad Q_{31} = \frac{K}{25}, \quad Q_{41} = \frac{K}{40}, \quad Q_{51} = \frac{K}{50}.$$

To determine K , we note that (why?)

$$\boxed{Q_{21} + Q_{31} + Q_{41} + Q_{51} = 1.}$$

Hence, we have

$$K(0.2 + 0.25 + 0.025 + 0.02) = 1 \quad \text{or} \quad K = 7.4074.$$

²You may consider $Q_{ij} \propto \frac{1}{C_{ij}+1}$ if C_{ij} is too close to 0

- We can then compute the remaining probabilities similarly.
- We have a new system as follows:

$$\begin{bmatrix} P_1(t+1) \\ P_2(t+1) \\ P_3(t+1) \\ P_4(t+1) \\ P_5(t+1) \end{bmatrix} = \begin{bmatrix} 0.0000 & 0.4000 & 0.3077 & 0.1515 & 0.1379 \\ 0.3704 & 0.0000 & 0.3846 & 0.0000 & 0.1724 \\ 0.2963 & 0.4000 & 0.0000 & 0.2424 & 0.0000 \\ 0.1852 & 0.0000 & 0.3077 & 0.0000 & 0.6897 \\ 0.1481 & 0.2000 & 0.0000 & 0.6061 & 0.0000 \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ P_4(t) \\ P_5(t) \end{bmatrix}.$$

In this case, we have

$$\mathbf{P}(7) = [0.1971 \ 0.1557 \ 0.2104 \ 0.1890 \ 0.2477]^T.$$

and $P_1(7) = 0.1971$.

- By solving the linear system of equations similar to Eq. (2.2), we can show that

$$\lim_{t \rightarrow \infty} \mathbf{P}(t) = [0.1929 \ 0.1786 \ 0.1857 \ 0.2357 \ 0.2071]^T.$$

2.4 A Counting Problem

- How many different ways of adding up 1's and 2's to 14?

$$\begin{array}{ll} 1 = 1 & T_1 = 1 \\ 2 = 1 + 1, 2 & T_2 = 2 \\ 3 = 1 + 1 + 1, 1 + 2, 2 + 1 & T_3 = 3 \\ 4 = 1 + 1 + 1 + 1, 1 + 1 + 2, 1 + 2 + 1, 2 + 1 + 1, 2 + 2 & T_4 = 5 \\ \vdots & \vdots \end{array}$$

- We let T_j be the number of ways of adding 1's and 2's to j .
- We may have the following relation (**Why?**):

$$\boxed{T_j = T_{j-1} + T_{j-2}.}$$

- Then you can compute T_{14} as you know T_1 and T_2 (Exercise).

$$T(1) = 1 \quad \text{and} \quad T(2) = 2,$$

For $j = 3 : 14$,

$$T(j) = T(j - 1) + T(j - 2);$$

end;

Output $T(14)$;

2.5 Some Mathematical Treatment for Second-order Homogeneous Difference Equations

- The previous problems are **second-order homogeneous difference equation** of the following form:

$$\boxed{a_2 T_{j+2} + a_1 T_{j+1} + a_0 T_j = 0, \quad j = 0, 1, 2, \dots,} \quad (2.3)$$

- We seek for solution of the form:

$$\boxed{T_j = z^j.}$$

If we substitute it into Eq. (2.3), we get the **Euler equation**:

$$\boxed{f(z) = a_2 z^2 + a_1 z + a_0 = 0.}$$

- The roots are given by

$$z_1 = \frac{-a_1 - \sqrt{a_1^2 - 4a_0a_2}}{2a_2}$$

and

$$z_2 = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_2}.$$

Case 1: Suppose that $z_1 \neq z_2$ and both are real numbers.

Then both

$$T_j = c_1 \cdot z_1^j \quad \text{and} \quad T_j = c_2 \cdot z_2^j$$

are solutions to Eq. (2.3) for any real numbers c_1 and c_2 .

- We can also check that their sum

$$T_j = c_1 \cdot z_1^j + c_2 \cdot z_2^j, \quad j = 0, 1, \dots, \tag{2.4}$$

is also a solution. Therefore it is the **general solution** to Eq. (2.3).

Case 2: If $z_1 = z_2$ (repeated root), then apart from the solution

$$T_j = c_1 \cdot z_1^j$$

another solution will be

$$T_j = c_2 \cdot j \cdot z_1^j.$$

Because when we substitute $T_j = j \cdot z_1^j$ into Eq. (2.3), we get

$$z_1^j \cdot (j \cdot (a_2 \cdot z_1^2 + a_1 \cdot z_1 + a_0) + (2 \cdot a_2 \cdot z_1^2 + a_1 \cdot z_1))$$

and

$$f(z_1) = a_2 \cdot z_1^2 + a_1 \cdot z_1 + a_0 = 0$$

and

$$z_1 \cdot (2 \cdot a_2 \cdot z_1 + a_1) = z_1 f'(z_1) = 0.$$

• Therefore the general solution is of the form:

$$T_j = z_1^j \cdot (c_1 + c_2 j), \quad j = 0, 1, \dots, \tag{2.5}$$

Case 3: Complex Roots. In this case, we consider the difference equation of the following form (where $a, b \in \mathbb{R}$):

$$\boxed{T_{j+2} - 2aT_{j+1} + (a^2 + b^2)T_j = 0.}$$

• Then $a + bi$ and $a - bi$ form a pair of complex (conjugate) roots of the equation:

$$x^2 - 2ax + (a^2 + b^2) = 0.$$

Then we have the solutions in their **polar form**:

$$a + bi = \rho(\cos \theta + i \sin \theta)$$

and

$$a - bi = \rho(\cos \theta - i \sin \theta)$$

where

$$\rho = \sqrt{a^2 + b^2} \quad \text{and} \quad \tan \theta = \left| \frac{b}{a} \right|. \quad (2.6)$$

- By **de Moivre's Theorem** we have for $j \in \mathbb{N}$

$$(a + bi)^j = \rho^j (\cos \theta + i \sin \theta)^j = \rho^j (\cos j\theta + i \sin j\theta)$$

and

$$(a - bi)^j = \rho^j (\cos \theta - i \sin \theta)^j = \rho^j (\cos j\theta - i \sin j\theta)$$

which can be proved by using mathematical induction on n (exercise).

- Thus

$$\begin{aligned} A_1(a + bi)^j + A_2(a - bi)^j &= A_1\rho^j(\cos j\theta + i \sin j\theta) + A_2\rho^j(\cos j\theta - i \sin j\theta) \\ &= (A_1 + A_2)\rho^j \cos j\theta + i(A_1 - A_2)\rho^j \sin j\theta. \end{aligned}$$

- Finally, in all the cases, the parameters (unknowns) can be solved if we are given extra information such as T_0 and T_1 , the initial conditions.

Exercise: $T_{j+2} - T_{j+1} + T_j = 0$ with $T_0 = 0$ and $T_1 = 1$.

2.6 A Rabbit Population Problem

- One of the early examples of a recursively defined sequence arises in the writings of **Fibonacci**, who was the greatest European Mathematician in the Middle Ages.
- In 1202, Fibonacci posed the following problem about the number of rabbits in a closed environment. He assumed that a single pair of rabbits is born at the beginning of a year. We then assume the following conditions:
 - (i) Rabbit pairs are **not fertile** during their first month of life but thereafter give birth to one new male/female pair at the end of every month.
 - (ii) No death occurs during the year
- The question is: How many rabbits will there be at the end of the year?

- Let T_j be the number of rabbit pairs alive at the end of month j . $T_0 = 1$, the initial number of rabbit pairs. We also have $T_1 = 1$
- The number of rabbit pairs at the end of month j is equal to the sum of **the number of rabbit pairs at the end of month $j - 1$** and **the number of rabbit pairs born at the end of month j** .
- **The number of rabbit pairs born at the end of month j is equal to the number of rabbit pairs at the end of month $j - 2$.**
- The recurrence relation ($j = 2, 3, \dots$) is then given by

$$T_j = T_{j-1} + T_{j-2} \quad \text{with} \quad T_0 = T_1 = 1.$$

Therefore the answer for the Fibonacci's question is $T_{12} = 233$. How?

- In fact, using our techniques in previous section, we can obtain (Exercise):

$$T_j = \frac{\sqrt{5} + 1}{2\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^j + \frac{\sqrt{5} - 1}{2\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^j.$$

2.7 An Eigenvalue Problem

- To obtain the eigenvalues and eigenvectors of a special $n \times n$ matrix:

$$A_n = \begin{bmatrix} 2 & -1 & & & & 0 \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 \\ 0 & & & & -1 & 2 \end{bmatrix}$$

- We consider the relation $A_n \mathbf{v} = \lambda \mathbf{v}$, then we have

$$\begin{cases} (2 - \lambda)v_1 - v_2 & = 0 \\ -v_1 + (2 - \lambda)v_2 - v_3 & = 0 \\ \vdots & \vdots \\ -v_{j-1} + (2 - \lambda)v_j - v_{j+1} & = 0 \\ \vdots & \vdots \\ -v_{n-1} + (2 - \lambda)v_n & = 0 \end{cases}$$

where $\mathbf{v} = [v_1, v_2, \dots, v_n]^T \neq \mathbf{0}$ is **an eigenvector**.

- We define $v_0 = v_{n+1} = 0$ then we have the difference equations:

$$-v_{j-1} + (2 - \lambda)v_j - v_{j+1} = 0 \quad \text{for } j = 1, 2, \dots, n.$$

or

$$\boxed{v_{j+1} - (2 - \lambda)v_j + v_{j-1} = 0 \quad \text{for } j = 1, 2, \dots, n.}$$

- Using the result we learned, the solution is of the form:

$$\boxed{v_j = Bm_1^j + Cm_2^j}$$

where B and C are two constants and m_1 and m_2 are roots of the quadratic equation

$$\boxed{m^2 - (2 - \lambda)m + 1 = 0} \tag{2.7}$$

- We remark that $m_1 \neq m_2$. Because if this is the case, we have the solution form:

$$v_j = (B + Cj)m_1^j$$

but $v_0 = v_{n+1} = 0$ implies that

$$v_0 = 0 = B \quad \text{and} \quad v_{n+1} = 0 = C(n + 1)m_1^{n+1}$$

and therefore $B = C = 0$ as $m_1 \neq 0$.

Now we have

$$v_0 = 0 = B + C$$

and

$$v_{n+1} = 0 = Bm_1^{n+1} + Cm_2^{n+1}.$$

Thus we have

$$B = -C \quad \text{and} \quad \left(\frac{m_1}{m_2}\right)^{n+1} = 1 = e^{2\pi i}$$

where $i = \sqrt{-1}$. Thus

$$\frac{m_1}{m_2} = e^{\frac{2s\pi i}{n+1}} \quad \text{for } s = 1, 2, \dots, n.$$

• By Eq. (2.7), we have

$$m_1 m_2 = 1.$$

Here we have

$$m_1 = e^{\frac{s\pi i}{n+1}} \quad \text{and} \quad m_2 = e^{-\frac{s\pi i}{n+1}}$$

- By Eq. (2.7) again, we have

$$m_1 + m_2 = 2 - \lambda$$

and thus we have for $s = 1, 2, \dots, n$,

$$\begin{aligned} \lambda_s &= 2 - \left(e^{\frac{s\pi i}{n+1}} + e^{\frac{-s\pi i}{n+1}} \right) \\ &= 2 - 2 \cos\left(\frac{s\pi}{n+1}\right) \quad (\cos(2\theta) = 1 - 2 \sin^2(\theta)) \\ &= 2 - 2\left(1 - 2 \sin^2\left(\frac{s\pi}{2(n+1)}\right)\right) \\ &= 4 \sin^2\left(\frac{s\pi}{2(n+1)}\right). \end{aligned}$$

Finally we note that for λ_s , we have $j = 1, 2, \dots, n$

$$v_j = Bm_1^j + Cm_2^j = B\left(e^{\frac{s\pi i}{n+1}} - e^{\frac{-s\pi i}{n+1}}\right) = 2iB \sin\left(\frac{j s \pi}{n+1}\right).$$

- As $2iB$ is just a constant, we have

$$\mathbf{v}_s = \left[\sin\left(\frac{s\pi}{n+1}\right), \sin\left(\frac{2s\pi}{n+1}\right), \dots, \sin\left(\frac{sn\pi}{n+1}\right) \right]^T.$$



Figure 3.1: St Augustine and Monica by Ary Scheffer (1846). Taken from Wikipedia.

3 Probability Theory, Poisson Process and Inventory Systems

3.1 The Science of Probability: A Historical Note

- With the advent of Christianity, the concept of **random events** developed by philosophers was rejected in the early time.
- According to **St. Augustine** (354-430), nothing occurred by chance, everything being minutely controlled by the will of God. If events appear to occur at **random**, then it is because of our ignorance and not in the nature of events. One should only seek for the will of God instead of looking at **patterns** of behavior in aggregates of events. ³

³Poker faces: the life and work of professional card players by David M. Hayano, UCP Press, 1982.

- The amazing contents and applications of probability theory owes its origin to a question on gambling (**game**).
- The question was raised by **Chevalier de Mere** (1607-1684) on his problem of throwing a die. He had a title Chevalier (Knight) and educated at Mere. The problem was solved by **Pascal**.

The Problem:

- De Mere made considerable money over the years in betting **double odds** on rolling at **least one “6”** in **4** throws of a fair die (six faces).
- He then thought that the same should occur for betting on at **least one double-six in 24 throws of two fair dice** (This is their ancient believes). It turned out that it did not work well.
- Why? In 1654, he challenged his friends Pascal and Fermat for reasons.



Figure 3.2: Pascal (1623-1662) (Left). Fermat (1601-1665) (Right). Taken from Wikipedia.

- The probability of getting no “6” in four independent throws of a fair die:

$$(5/6) \times (5/6) \times (5/6) \times (5/6) = \frac{625}{1296}.$$

- Therefore the probability of having **at least one** “6” in 4 throws will be equal to

$$1 - \frac{625}{1296} = \frac{671}{1296} = 0.5177 > 0.5000.$$

- This explained why de Mere got a good amount of money on double odds on his bet.
- This is **NOT** a fair game, the player has advantage over the house.

- The probability of getting no double “6” in the throw of two fair dice is

$$1 - \left(\frac{1}{6} \times \frac{1}{6}\right) = \frac{35}{36}.$$

- The probability of getting no double “6” in “24” independent throws is

$$\left(\frac{35}{36}\right)^{24}.$$

- Therefore, the probability of having at least one double 6 in 24 throws is equal to

$$1 - \left(\frac{35}{36}\right)^{24} = 0.4914 < 0.5.$$

- This explained why de Mere did not get a good amount of money on double odds on this bet.
- Again this is **NOT** a fair game, the house has advantage over the player.

3.2 Poisson Process and Inventory Systems

- Everyday we encounter with a lot of random events. The **waiting time** for a bus, the **service time** at a counter, the **lifetime** of a machine, the **processing time** for a product etc.

The **waiting time for the occurrence of an event** can be modeled by a **continuous probability distribution**.

- On the other hand, there are other events such as: the **number of buses** passes through a junction in one hour, the **number of customers** served per day, the **number of broken machines** handled in a year and the **number of products produced** per month.

The **number of events** occurs in a **fixed period** can be modeled by a **discrete probability distribution**.

- Here we propose to employ the **Exponential distribution** (continuous case)

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0, \quad \lambda > 0.$$

and the **Poisson distribution** (discrete case)

$$f(n, t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad \lambda > 0, \quad t \geq 0, \quad n = 0, 1, \dots$$

to model the above random events.

- These two distributions are linked up by the **Poisson process**. We will also discuss the relationship in detail.

3.3 Exponential Distribution and Poisson Distribution

3.3.1 Exponential Distribution

Definition 3.1. A continuous random variable X is said to follow an **exponential distribution** with parameter λ , if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Proposition 3.1. *If a random variable X follows the exponential distribution with parameter λ then*

$$E(X) = \lambda^{-1}$$

and

$$\text{Var}(X) = \lambda^{-2}$$

and its moment generating function is given by

$$\lambda(\lambda - t)^{-1}.$$

Proof. First by definition, we have

$$E(X) = \int_0^{\infty} x f(x) dx = \frac{1}{\lambda}$$

and

$$E(X^2) = \int_0^{\infty} x^2 f(x) dx = \frac{2}{\lambda^2}.$$

Therefore, we have

$$Var(X) = E(X^2) - E(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Finally, the moment generating function is then given by

$$M(t) = E(e^{tX}) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}.$$

We note that $M'(0) = \lambda^{-1} = E(X)$ and $M''(0) = 2\lambda^{-2} = E(X^2)$. □

Remark 3.1. We note that

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx.$$

There is a one-to-one relation between the **Probability Density Function** (PDF) and its **Moment Generating Function** (MGF).

Definition 3.2. A probability distribution (let say a non-negative random variable X) is said to have the Markov property if for any two non-negative values t and x we have

$$P\{X > x + t | X > t\} = P\{X > x\}.$$

Proposition 3.2. *The exponential distribution has the Markov property.*

Proof. This follows from

$$\begin{aligned} P\{X > t + x | X > t\} &= \frac{e^{-\mu(t+x)}}{e^{-\mu t}} \\ &= e^{-\mu x} \\ &= P\{X > x\}. \end{aligned} \tag{3.1}$$

□

- In a number of applications, observation has shown that the exponential distribution can be a good description of service time distribution (which is therefore called **exponential service time**).
- Exponentially distributed time has the nice feature that by the Markov property (3.1), the distribution of the remaining holding time after a customer has been served for any length of time $t > 0$ is the same as that initially at $t = 0$.

Here are some more properties of the exponential distribution.

Proposition 3.3. *If X_1 (the waiting time for Bus 1) and X_2 (the waiting time for Bus 2) are **independent non-negative** random variables with probability density functions $f_1(t)$ and $f_2(t)$. If they are both **exponentially distributed** random variables with means μ_i^{-1} , then the probability that X_2 exceeds X_1 is $\mu_1/(\mu_1 + \mu_2)$.*

Proof.

$$P\{X_2 > X_1\} = \int_0^\infty \int_s^\infty f_1(s)f_2(t)dt ds = \int_0^\infty \int_0^t f_1(s)f_2(t)ds dt$$

i.e., we integrate the joint density function over the region

$$\Omega = \{(s, t) \in \mathbb{R}^2 | t > s\}.$$

If X_1 and X_2 are exponential with mean μ_1^{-1} and μ_2^{-1} , then the above integral becomes

$$\int_0^\infty \int_s^\infty \mu_1 e^{-\mu_1 s} \mu_2 e^{-\mu_2 t} dt ds = \int_0^\infty \mu_1 e^{-\mu_1 s} e^{-\mu_2 s} ds = \frac{\mu_1}{\mu_1 + \mu_2}. \quad (3.2)$$

□

Proposition 3.4. *Suppose that X_1, X_2, \dots, X_n are **independent, identical, exponential** random variables with mean μ^{-1} , and consider the corresponding order statistics*

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

Then $X_{(1)}$ is again exponentially distributed with mean $\frac{1}{n}$ times the mean of the original random variables.

Proof. We observe that

$$X_{(1)} = \min(X_1, X_2, \dots, X_n).$$

- We note $X_{(1)} > x$ **if and only if** all $X_i > x$ ($i = 1, 2, \dots, n$).

Hence we have

$$\begin{aligned} P\{X_{(1)} > x\} &= P\{X_1 > x\}P\{X_2 > x\} \cdots P\{X_n > x\} \\ &= (e^{-\mu x})^n \\ &= e^{-n\mu x}. \end{aligned}$$

□

Proposition 3.5. *A random variable X is exponentially distributed **if and only if***

$$P\{X < t + h | X > t\} = \mu h + o(h) \text{ as } h \rightarrow 0.$$

Here $o(h)$ is a function of h such that

$$\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0.$$

Proof. Suppose X has an exponential distribution,

$$\begin{aligned} P\{X < t + h | X > t\} &= 1 - e^{-\mu h} && \text{(Markov property)} \\ \text{(Apply Taylor's series, see Remark 3.2)} &= 1 - (1 - \mu h + o(h)) && (3.3) \\ &= \mu h + o(h) \text{ as } h \rightarrow 0. \end{aligned}$$

• **Conversely**, suppose that

$$P\{X < t + h | X > t\} = \mu h + o(h) \text{ as } h \rightarrow 0$$

then we have

$$P\{X > t + h | X > t\} = 1 - \mu h + o(h).$$

Using

$$P\{X > t + h | X > t\} = \frac{P\{X > t + h\}}{P\{X > t\}},$$

re-arranging the terms and let $h \rightarrow 0$, one can obtain the differential equation (see Remark 3.3)

$$\frac{d}{dt}P\{X > t\} = -\mu P\{X > t\},$$

which has the unique solution

$$P\{X > t\} = e^{-\mu t}$$

satisfying the initial condition

$$P\{X > 0\} = 1,$$

i.e. X follows the **Exponential distribution**. □

Remark 3.2. For $0 < h < 1$ we have

$$\begin{aligned}
 1 - e^{-\mu h} &= 1 - \left(1 - \frac{\mu h}{1!} + \frac{(\mu h)^2}{2!} - \frac{(\mu h)^3}{3!} + \dots\right) \\
 &= \frac{\mu h}{1!} - \frac{(\mu h)^2}{2!} + \frac{(\mu h)^3}{3!} - \dots \\
 &\equiv \mu h + g(h).
 \end{aligned}$$

Then, one can see that

$$\begin{aligned}
 \left| \frac{g(h)}{h} \right| &= \left| \left(-\frac{(\mu)^2 h}{2!} + \frac{(\mu)^3 h^2}{3!} - \dots \right) \right| \\
 &\leq h \left(\frac{(\mu)^2}{2!} + \frac{(\mu)^3}{3!} + \dots \right) \\
 &= h e^{\mu}.
 \end{aligned}$$

Therefore, we have

$$\lim_{h \rightarrow 0} \frac{g(h)}{h} = 0 \quad \text{and hence} \quad g(h) = o(h).$$

Remark 3.3. Given

$$P\{X < t + h | X > t\} = \mu h + o(h) \quad \text{as } h \rightarrow 0$$

then we have

$$P\{X > t + h | X > t\} = 1 - \mu h + o(h).$$

Using

$$P\{X > t + h | X > t\} = \frac{P\{X > t + h\}}{P\{X > t\}}$$

we have

$$P(X > t + h) = P(X > t) - \mu h P(X > t) + o(h) P(X > t).$$

Let $F(t) = P(X > t)$ then we have

$$\begin{aligned} F(t+h) &= F(t) - \mu h F(t) + o(h)F(t) \\ &= F(t) - \mu h F(t) + o(h). \end{aligned}$$

• Thus

$$\left| \frac{F(t+h) - F(t)}{h} - (-\mu F(t)) \right| = \left| \frac{o(h)}{h} \right|.$$

Hence

$$\lim_{h \rightarrow 0} \left| \frac{F(t+h) - F(t)}{h} - (-\mu F(t)) \right| = \lim_{h \rightarrow 0} \left| \frac{o(h)}{h} \right| = 0.$$

• This means

$$F'(t) = -\mu F(t).$$

And $F(t)$ can be solved easily with $F(0) = 1$.

• Solving the differential equation, we have

$$F(t) = e^{-\mu t}.$$

3.3.2 Poisson Distribution

Poisson distribution is a **discrete distribution** given by

$$P_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \quad j = 0, 1, 2, \dots$$

Proposition 3.6. *If a random variable X follows the Poisson distribution then*

$$E(X) = \lambda t$$

and

$$\text{Var}(X) = \lambda t$$

and its moment generating function is

$$e^{-\lambda t(1-e^z)}.$$

Proof. We have

$$\begin{aligned} E(X) &= \sum_{j=0}^{\infty} j \left(\frac{(\lambda t)^j}{j!} e^{-\lambda t} \right) \\ &= \sum_{j=1}^{\infty} (\lambda t) \frac{(\lambda t)^{j-1}}{(j-1)!} e^{-\lambda t} = \lambda t. \end{aligned}$$

• Since

$$\begin{aligned} E(X(X-1)) &= \sum_{j=1}^{\infty} j(j-1) \left(\frac{(\lambda t)^j}{j!} e^{-\lambda t} \right) \\ &= \sum_{j=2}^{\infty} (\lambda t)^2 \frac{(\lambda t)^{j-2}}{(j-2)!} e^{-\lambda t} \\ &= (\lambda t)^2. \end{aligned}$$

We have

$$E(X^2) - E(X) = (\lambda t)^2.$$

- Hence, we have

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ &= (\lambda t)^2 + E(X) - E(X)^2 \\ &= \lambda t. \end{aligned}$$

- Finally, the moment generating function is given by

$$\begin{aligned} E(e^{Xz}) &= \sum_{j=0}^{\infty} e^{jz} \left(\frac{(\lambda t)^j}{j!} e^{-\lambda t} \right) \\ &= \sum_{j=0}^{\infty} \left(\frac{(\lambda t e^z)^j}{j!} e^{-\lambda t} \right) \\ &= e^{-\lambda t + \lambda t e^z} \\ &= e^{-\lambda t(1 - e^z)}. \end{aligned}$$

□

Proposition 3.7. *Sum of two **independent Poisson** variables is again a Poisson variable.*

Proof. Let

$$X = X_1 + X_2$$

where X_i is a Poisson random variable with mean $\lambda_i t$. Then X has the moment generating function

$$g(z) = e^{-\lambda_1 t(1-e^z)} e^{-\lambda_2 t(1-e^z)} = e^{-(\lambda_1 + \lambda_2)t(1-e^z)}.$$

This shows that the sum of two independent Poisson random variable with means $\lambda_1 t$ and $\lambda_2 t$ is itself a Poisson random variable with mean

$$(\lambda_1 + \lambda_2)t.$$

□

Remark 3.4. Let $g_i(z)$ be the **moment generating functions** of X_i . If X_1 and X_2 are independent, then the moment generating function of $X_1 + X_2$ is given by

$$E(e^{t(X_1+X_2)}) = E(e^{tX_1} \cdot e^{tX_2}) = E(e^{tX_1}) \cdot E(e^{tX_2}) = g_1(z) \cdot g_2(z).$$

3.4 Poisson Process

Definition 3.3. A process $P(t)$ is said to be a Poisson process if at any epoch t ,

$$P\{\text{one occurrence during } (t, t+h)\} = \lambda h + o(h) \quad \text{as } h \rightarrow 0$$

and

$$P\{\text{two or more occurrence during } (t, t+h)\} = o(h) \quad \text{as } h \rightarrow 0.$$

Remark 3.5. The following are some remarks on the notation $o(t)$.

(a) We recall that $o(h)$ is a function of h such that

$$\boxed{\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0.}$$

(b) Possible examples of $o(h)$ are $o(h) = h^2$ and $o(h) = h \sin(h)$.

(c) However, $o(h)$ cannot take the form \sqrt{h} or $h \log(h)$.

Poisson process, **Poisson distribution** and **exponential distribution** are related as follows:

Proposition 3.8. *Suppose in a certain process, we let T_i ($i = 1, 2, 3, \dots$) be the epoch of the i^{th} occurrence.*

Let $A_i = T_i - T_{i-1}$ ($i = 1, 2, 3, \dots$); $T_0 =$ epoch that we start to count the number of occurrences.

Let $X(t) =$ number of occurrences in a time interval of length t . Then the following statements are equivalent.

- (a) *The process is **Poisson** (with mean rate λ).*
- (b) *$X(t)$ is a **Poisson** random variable with parameter λt , i.e.*

$$P\{X(t) = j\} = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \quad j = 0, 1, 2, \dots .$$

- (c) *A_i 's are mutually independent identically distributed exponential random variables with mean λ^{-1} , i.e.*

$$P\{A_i \leq t\} = 1 - e^{-\lambda t}, \quad i = 1, 2, \dots .$$

(a) implies (b)

- Given a Poisson process, with mean rate λ , we want to find the number of occurrences in the time interval $[0, t]$ and the probability density of the time between two successive occurrences.
- Divide the time interval into n equal parts and each is of length t/n .

In each of the sub-interval, the probability that it contains an occurrence is given roughly by $\lambda h = \lambda t/n$. Therefore, the probability that one get k occurrences is given by the **Binomial distribution**

$$\binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} = \underbrace{\frac{n \cdots (n - k + 1)}{n^k}}_{\rightarrow 1} \cdot \frac{(\lambda t)^k}{k!} \cdot \underbrace{\left(1 - \frac{\lambda t}{n}\right)^{n-k}}_{\rightarrow e^{-\lambda t}}.$$

- By letting n goes to infinity, from the well-known result, we have the Poisson distribution: $\frac{e^{-\lambda t} (\lambda t)^k}{k!}$.

(b) implies (a)

We note that

$$\begin{aligned}P(k = 0) &= e^{-\lambda t} \\ &= 1 - \lambda t + o(t) \quad (\text{by Remark 3.2})\end{aligned}$$

and

$$\begin{aligned}P(k = 1) &= e^{-\lambda t} \lambda t \\ &= \lambda t \left(1 - \lambda t + \frac{(\lambda t)^2}{2!} + \cdots + \right) \\ &= \lambda t + \left(-(\lambda t)^2 + \frac{(\lambda t)^3}{2!} + \cdots + \right) \\ &= \lambda t + o(t).\end{aligned}$$

Therefore we have

$$\begin{aligned}P(k > 1) &= 1 - P(k = 0) - P(k = 1) \\ &= o(t).\end{aligned}$$

(a) and (b) implies (c)

Let $f(t)$ be the probability density of the inter-occurrence time T .

Then, by definition, we have

$$\begin{aligned} P(t \leq T \leq t + \delta t) &= f(t)\delta t \\ &= P(\text{no occurrence in } (0, t)) \times P(1 \text{ occurrence in } (t, t + \delta t)). \end{aligned}$$

Now, we have

$$P(\text{no occurrence in } (0, t)) = e^{-\lambda t}$$

and

$$P(\text{one occurrence in } (t, t + \delta t)) = \lambda \delta t.$$

Hence we have

$$f(t)\delta t = \lambda e^{-\lambda t} \delta t$$

and

$$f(t) = \lambda e^{-\lambda t}.$$

(c) implies (b)

We will prove by using **mathematical induction** that the PDF of the random variable $A(n) = A_1 + A_2 + \dots + A_n$ is given by the **Erlangian distribution**:

$$g_n(t) = \frac{\lambda(\lambda t)^{n-1}e^{-\lambda t}}{(n-1)!}.$$

For $n = 1$, it is clear the proposition is true and we assume the statement holds for some positive integer n . We then look for the PDF of

$$Z = A_n + \mathbf{A}.$$

We have

$$\begin{aligned} P(Z \leq t) &= \int_0^t \int_0^{t-x} g_n(x) \lambda e^{-s\lambda} ds dx \\ &= \int_0^t g_n(x) (1 - e^{-\lambda(t-x)}) dx \\ &= \int_0^t g_n(x) dx - e^{-\lambda t} \int_0^t \frac{\lambda(\lambda x)^{n-1}}{(n-1)!} dx \\ &= \int_0^t g_n(x) dx - e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \end{aligned}$$

- Taking the derivative of the R.H.S. with respect to t , we obtain the PDF of Z as follows:

$$g_{n+1}(t) = \frac{\lambda(\lambda t)^n e^{-\lambda t}}{n!}.$$

We have the following holds for $n = 1, 2, \dots$,

$$\begin{aligned} P(A(n) \leq t) &= \int_0^t \frac{\lambda(\lambda x)^{n-1} e^{-\lambda x}}{(n-1)!} dx \\ &= 1 - \sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad (\text{apply integration by parts}) \\ &= 1 - \sum_{k=0}^{n-1} P(X(t) = k). \end{aligned}$$

- Here $P(A(n) \leq t)$ is the probability that the waiting time for the n th arrivals is less than or equal to t .
- This is also the probability that there are n or more arrivals observed in the interval $[0, t]$. Thus $X(t)$ follows the Poisson distribution with mean rate λt .

Remark 3.6. Poisson process provides rather good approximation for modeling many random processes such as the arrival of customers and calls. From the proposition above, a process is Poisson (with coefficient λ) if and only if the **inter-arrival times** (the lengths of time between successive customer arrivals) are mutually independent exponentially distributed random variables with mean λ^{-1} .

- We note that if the mean arrival rate is λ then the mean inter-arrival time is λ^{-1} . By the Markov property of exponential random variables, the distribution of lengths of time from an arbitrarily chosen epoch to the next arrival (called the **next-arrival times**) is the same as the distribution of inter-arrival times.

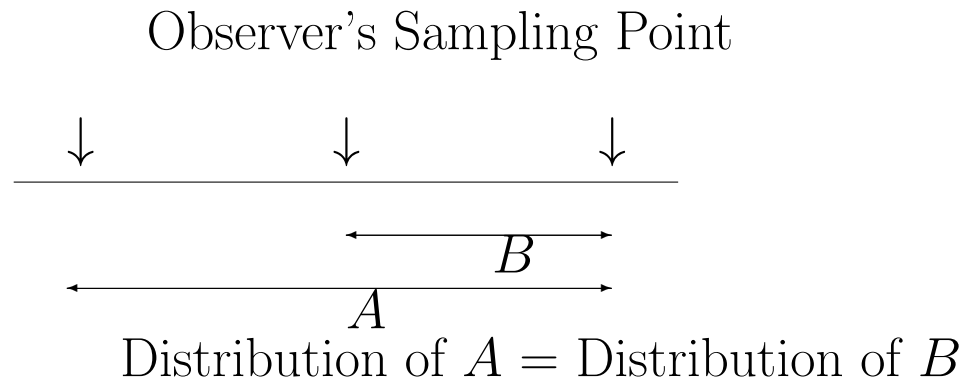


Figure 3.3. Arrival of Customers.

- This nice property much simplifies the mathematical analysis of many situations.

- For a Poisson process,

$$P\{\text{exactly one occurrence in } (t, t + h)\} = \lambda h e^{-\lambda h}.$$

Now for a fixed t and any x in $(0, t)$,

$$\begin{aligned} & P\{\text{the epoch of occurrence is in } (0, x) \mid \text{exactly one occurrence in } (0, t)\} \\ &= \frac{P\{\text{exactly one occurrence in } (0, x), \text{ and no occurrence in } (x, t)\}}{P\{\text{exactly one occurrence in } (0, t)\}} \\ &= \frac{\lambda x e^{-\lambda x} \times e^{-\lambda(t-x)}}{\lambda t e^{-\lambda t}} \\ &= \frac{x}{t}, \end{aligned}$$

which is a uniform distribution.

- This means that if we know there is exactly one occurrence in $(0, t)$ then the epoch of that occurrence is equally likely throughout $(0, t)$. In this sense we say that a **Poisson process is random**.

3.5 Law of Large Numbers

In this section, we first introduce without prove the famous **law of large numbers** which states that the long-run average of a sequence of independent and identically distributed random variables will converge to its mean. We then prove the **weak law of large numbers**. We state without proving the **strong law of large numbers**.

- Instead we shall state and give a proof for the **weak law of large numbers**.
- At the same time we also introduce two important inequalities for random variables: **Markov inequality** and **Chebyshev inequality**.
- By making use of the Chebyshev inequality, one can prove the **weak law of large numbers**.
- Let us begin our proof with the Markov inequality.

Proposition 3.9. *If a random variable takes X only **non-negative values**, then for any $a > 0$, we have*

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

*This is known as the **Markov inequality**.*

Proof. We give a proof for the case when X is a continuous random variable. The case for discrete random variable is similar and therefore omitted. Let $f(x)$ be the probability density function.

$$\begin{aligned} E(X) &= \int_0^{\infty} x f(x) dx \\ &= \int_0^a x f(x) dx + \int_a^{\infty} x f(x) dx \\ &\geq \int_a^{\infty} x f(x) dx \\ &\geq \int_a^{\infty} a f(x) dx \quad (\text{because } x f(x) \geq a f(x), \text{ for } x \geq a) \\ &= a \int_a^{\infty} f(x) dx = a P(X \geq a). \end{aligned}$$

□

- Using the Markov inequality, one can also prove the Chebyshev's inequality.

Proposition 3.10. *If a random variable X has mean μ and variance σ^2 then for any $k > 0$, we have*

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

This is known as the **Chebyshev's inequality**.

Proof. Let $(X - \mu)^2/\sigma^2$ be a non-negative random variable whose mean is

$$E\left[\frac{(X - \mu)^2}{\sigma^2}\right] = \frac{E[(X - \mu)^2]}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1.$$

Let $a = k^2$ and apply the Markov inequality we have

$$P\left(\frac{(X - \mu)^2}{\sigma^2} \geq k^2\right) \leq \frac{1}{k^2}.$$

Hence we get

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

□

Proposition 3.11. *Let X_1, X_2, \dots, X_n be a sequence of independent and identical distributed random variables having mean μ and finite variance σ^2 . Then for any $e > 0$ we have*

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \geq e \right) = 0.$$

*This is the famous **Weak Law of Large Numbers**.*

Proof. Let

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

we have

$$E(\bar{X}) = \frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n} = \frac{\mu + \mu + \dots + \mu}{n} = \mu$$

and

$$\text{Var}(\bar{X}) = \frac{\text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)}{n^2} = \frac{\sigma^2 + \sigma^2 + \dots + \sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

By Chebyshev's inequality for any positive k we have

$$P(|\bar{X} - \mu| \geq \frac{k\sigma}{\sqrt{n}}) \leq \frac{1}{k^2}.$$

In particular we let $k = e\sqrt{n}/\sigma$ and get

$$P(|\bar{X} - \mu| \geq e) \leq \frac{\sigma^2}{ne^2}.$$

Therefore for any positive e we have

$$\lim_{n \rightarrow \infty} P(|\bar{X} - \mu| \geq e) = 0.$$

Here let us also state without prove the *strong law of large numbers*. A generalization of the weak law of large numbers is that, with probability one,

$$\lim_{n \rightarrow \infty} \bar{X} = \mu.$$

This means that the long-run average of a sequence of independent and identically distributed random variables will converge to its mean. \square

Proposition 3.12. *Let X_1, X_2, \dots, X_n be a sequence of independent and identical distributed random variables having mean μ and finite variance σ^2 . Then we have*

$$P\left(\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} = \mu\right) = 1.$$

*This is the famous **Strong Law of Large Numbers**.*

3.6 Introduction to Inventory Models

- To meet demand on time and compete in the market, a company needs to keep suitable amount of stock in hand and order replenishment at a right time. Too much stock incurs extra inventory cost but shortage in stock will result in unsatisfied demands.
- The purpose of inventory theory is to determine rules or policies so as to minimize the system running cost (or maximize its profit) and meet the customer demand.
- To obtain “good” operational policy, one has to understand the process of **demand**, **flow of material** and the **cost of the facility** etc. Mathematical models, especially probabilistic models are good tools for the analysis of inventory process.
- The main objective here is to obtain optimal ordering policies for inventory systems such that the overall system running cost is minimized. We will begin with simple deterministic models and then complex probabilistic models.

The following are possible costs associated with an inventory system.

1. **Ordering and setup cost:** it includes the cost of paperwork and billing associated with an order. If the product is produced internally, this cost may also include the cost for setting up a machine in a production system and also the cost for labor.
2. **Purchasing cost:** it includes the cost of raw material and transportation cost.
3. **Holding cost:** this is the cost of holding an unit of inventory for one period. If the period is one year than it is the annual holding cost. This cost may also include the insurance cost due to the possibility of spoilage.
4. **Shortage cost:** when a demand cannot be met in time, a shortage is said to be occurred. There are two possible cases, either
 - (i) the customers accept delivery on a later date, this is called a **backlogged demand** or
 - (ii) the customers refuse to have the delivery on a later date and this is called a **lost sale**.

3.6.1 Deterministic (EOQ) Inventory Models

The followings are the basic assumptions of the deterministic inventory Models.

1. Repetitive ordering:

This means that the ordering decision is repetitive in the sense that it is repeated in a regular manner.

2. Constant demand:

Demand is assumed at a constant rate, let say D per year. Then the total demand for t years will be Dt .

3. Constant lead time:

The lead time for each order is a known constant L .

Definition 3.4. The lead time is the length of time between the instance **when an order is placed** and the instance at which **the order arrives**.

3.6.2 The Basic EOQ Model

The basic assumptions of a **Economic Order Quantity** (EOQ) Model are as follow.

1. The **demand** D per year (we assume the unit of time is one year) is a known constant,
2. The **ordering cost** is K when an order is placed,
3. The **size of an order** q is a constant to be determined,
4. The **lead time** L is assumed to be 0,
5. No shortage of demand is allowed,
6. The **holding cost** for one unit of inventory is h per year,
7. The **unit purchasing cost** is p .

Remark 3.7. We note that the orders arrive instantaneously, in an optimal policy one should never place an order when the inventory level is greater than zero or otherwise an unnecessary holding cost will be incurred. When the inventory level is zero, one must place an order to prevent a shortage from occurring. The inventory level of the system is given in Figure 1.5.

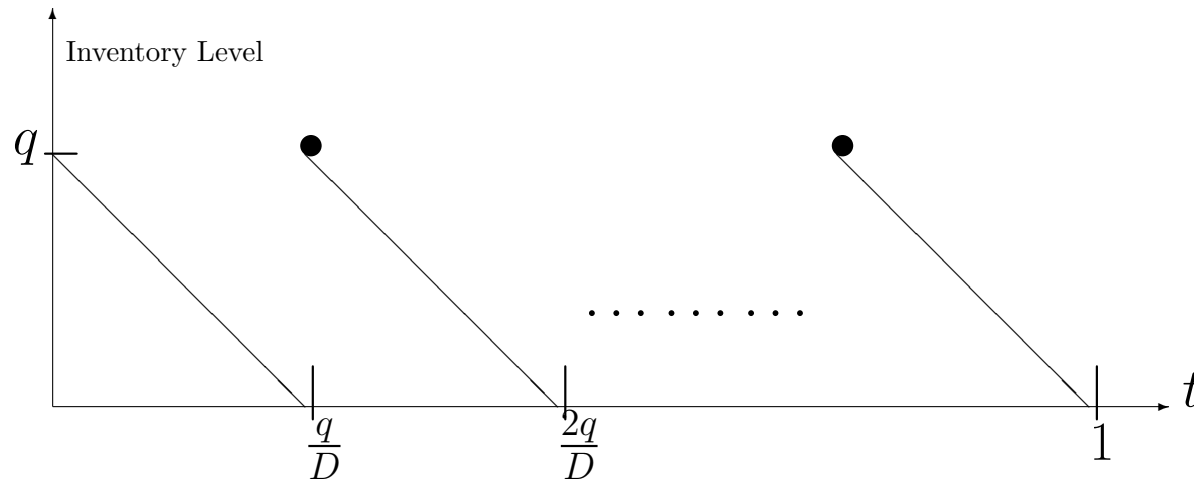


Figure 3.4. The Basic EOQ Model.

Definition 3.5. Any interval of time that begins with the arrival of an order and ends with the instant before the next order is received is called a cycle.

Remark 3.8. We note the **cycle** of this model is q/D .

- We are going to determine the optimal value of q such that the total annual running cost of the system is minimized.
- The annual running cost $c(q)$ can be written as follows:

cost of placing order + purchasing cost + holding cost.

(a) Since the size of each order is q units and the annual demand is D , one needs D/q orders per year. Thus the annual cost of placing order is given by

$$K\left(\frac{D}{q}\right).$$

(b) The annual purchasing cost is of course given by

$$pD.$$

(c) In each cycle, inventory level decreases from q down to 0 at a linear rate of D .

Thus the average inventory level is given by

$$\frac{\frac{1}{2} \cdot q/D \cdot q}{q/D} = \frac{q}{2}.$$

The annual holding cost is given by

$$h\left(\frac{q}{2}\right).$$

• From (a)-(c), the annual cost is given by

$$\boxed{c(q) = \frac{KD}{q} + pD + \frac{hq}{2}}. \quad (3.4)$$

- We note that

$$c'(q) = \frac{-KD}{q^2} + \frac{h}{2}.$$

The only critical point is obtained by solving $c'(q) = 0$ and We have

$$q^* = \sqrt{\frac{2KD}{h}}.$$

Since

$$c''(q) = \frac{2KD}{q^3} > 0 \quad \text{for } q > 0$$

the annual cost $c(q)$ is minimized when $q = q^*$. The optimal size of an order

$$q^* = \sqrt{\frac{2KD}{h}}$$

is called the **Economic Order Quantity** (EOQ).

- In this case the optimal annual cost is given by

$$c(K, D, h) = KD\sqrt{\frac{h}{2KD}} + pD + \frac{h}{2}\sqrt{\frac{2KD}{h}} = \sqrt{2hKD} + pD. \quad (3.5)$$

Remark 3.9. We give some remarks on the EOQ Model.

1. The EOQ does NOT depend on the unit purchasing price, because no matter what is the size of each order, one has to purchase D units per year.
2. From the formula, we note that EOQ increases when the annual demand D or the ordering cost K increases.
3. If the unit holding cost increases, then the EOQ decreases and the system holds less inventory.
4. We also note is that when EOQ is applied, the annual holding cost is given by

$$\frac{hq}{2} = \frac{h}{2} \sqrt{\frac{2KD}{h}} = \sqrt{\frac{hKD}{2}}$$

and the annual ordering cost is given by

$$\frac{KD}{q} = KD \sqrt{\frac{h}{2KD}} = \sqrt{\frac{hKD}{2}}.$$

Therefore the annual holding cost is equal to the annual ordering cost.

Example 3.1. A company uses 500 units of drug per year. The cost for the ordering of drug is 20, each unit of drug cost 100 and the holding cost for one unit of drug is 2 per year.

- (a) What is the EOQ and the optimal running cost?
- (b) How many order will be placed each year?
- (c) What is the cycle length?

We note that in this case $K = 20$, $h = 2$ and $D = 500$.

- (a) The EOQ will be

$$\sqrt{\frac{2KD}{h}} = \sqrt{\frac{2(20)500}{2}} = 100.$$

The optimal running cost is

$$\sqrt{2hKD} + pD = \sqrt{2(2)(20)(500)} + 100(500) = 50200.$$

- (b) The number of orders will be $500/100 = 5$.
- (c) The cycle length will be $1/5$ year.

Example 3.2. In the previous example, if the drug is rotten and cannot be used if it spends more than one month in the inventory. What is the EOQ, optimal running cost and the cycle length in this case?

- In previous example the cycle length is 2.4 month which is longer than one month. Thus the new cycle length is one month (1/12 year).

- The new EOQ is

$$D/12 = 500/12$$

and the optimal running cost is

$$\begin{aligned} & \frac{KD}{q} + pD + \frac{hq}{2} \\ &= \frac{20(500)}{500/12} + (100)500 + \frac{2(500/12)}{2} \\ &= 50281.7. \end{aligned}$$

3.6.3 When EOQ is Applicable

- One main assumption on the deterministic EOQ model is the constant demand.
- To determine if the assumption of constant demand is reasonable, suppose that the demand during the n periods of time D_1, D_2, \dots, D_n are observed. Let the mean of the observed demands be

$$\bar{D} = \frac{1}{n} \left(\sum_{i=1}^n D_i \right)$$

and the variance of the observed demands be

$$\text{Var}(D) = \frac{1}{n} \left(\sum_{i=1}^n D_i^2 \right) - \bar{D}^2.$$

- We define **coefficient of variation** as

$$\text{CV} = \frac{\sqrt{\text{Var}(D)}}{\bar{D}}.$$

Clearly if $\text{CV} = 0$, then the demand is constant. In general if

$$\text{CV} < 0.2$$

then the assumption of constant demand is reasonable.

3.6.4 Newsboy Model with Discrete Stochastic Demand

In the previous section we studied some inventory models where the demand is a known constant. In this section we will discuss inventory models with stochastic (uncertain) demand. In our discussion, stochastic (uncertain) means that the demand is no longer constant but follows certain known probability function. We begin our discussion with the newsboy model.

- A newsboy sells newspaper (**perishable**) every morning. The cost of each newspaper remains at the end of the day is C_o (**overage cost**) and the cost of each unsatisfied demand is C_u (**underage cost or shortage cost**).
- Suppose that the probability density function of the demand D is given by

$$\text{Prob}(D = i) = p_i \geq 0, \quad i = 0, 1, 2, \dots$$

- Our aim is to determine the best amount Q of newspaper to be ordered such that the expected cost is minimized.

- Consider the following newsboy problem. If the daily demand is given as follows:

x	0	1	2	3	4
$P(x)$	0.05	0.2	0.3	0.4	0.05

Table 3.1.

- Suppose the unit underage cost and the overage cost are the **same**, what is the optimal ordering quantity Q^* ?
- It is clear that the optimal order quantity Q^* should be in $\{0, 1, 2, 3, 4\}$.

Most likely $Q^* \neq 0$ and $Q^* \neq 4$.

It is possible that $Q^* = 1, 2, 3$.

Most likely Q^* should be 3 as $P(x = 3) = 0.4$, it has the highest probability.

Do you agree? You will find out the answer yourself later.

Proposition 3.13. Define the cumulative probability function of the demand D to be

$$F(Q) = \sum_{i=0}^Q p_i = \text{Prob}(D \leq Q)$$

Then the optimal value of Q satisfies

$$F(Q - 1) < \frac{C_u}{C_u + C_o} \leq F(Q).$$

Step 1: We first construct the expected cost for a given order size Q . Clearly we have the following TWO cases.

(i) If the demand $D < Q$, then the cost will be $(Q - D)C_o$ and

(ii) if the demand $D > Q$, then the cost will be $(D - Q)C_u$.

Therefore the expected cost is given by

$$E(Q) = \underbrace{C_o \sum_{i=0}^Q (Q - i)p_i}_{\text{Expected Overage Cost}} + \underbrace{C_u \sum_{i=Q+1}^{\infty} (i - Q)p_i}_{\text{Expected Shortage Cost}}.$$

We note that $F(Q)$ is an increasing function in Q , i.e.

$$F(Q + 1) \geq F(Q) \quad \text{for } Q = 0, 1, 2, \dots$$

and $F(Q) \rightarrow 1$ as $Q \rightarrow \infty$.

We are going to show that the optimal Q is the one which satisfies

$$F(Q - 1) < \frac{C_u}{C_u + C_o} \leq F(Q).$$

Step 2: From the definition of the expected cost $E(Q)$, we are going to show

$$E(Q) - E(Q + 1) = C_u - (C_o + C_u)F(Q) \tag{3.6}$$

and

$$E(Q) - E(Q - 1) = -C_u + (C_o + C_u)F(Q - 1). \tag{3.7}$$

To prove (3.6), we recall that the expected cost

$$E(Q) = C_o \sum_{i=0}^Q (Q - i)p_i + C_u \sum_{i=Q+1}^{\infty} (i - Q)p_i.$$

Now we have

$$\begin{aligned}
E(Q+1) &= C_o \sum_{i=0}^{Q+1} (Q+1-i)p_i + C_u \sum_{i=Q+2}^{\infty} (i-Q-1)p_i \\
&= [C_o \sum_{i=0}^{Q+1} (Q-i)p_i + C_o \sum_{i=0}^{Q+1} p_i] + [C_u \sum_{i=Q+2}^{\infty} (i-Q)p_i - C_u \sum_{i=Q+2}^{\infty} p_i] \\
&= [C_o \sum_{i=0}^Q (Q-i)p_i - C_o p_{Q+1} + C_o \sum_{i=0}^{Q+1} p_i] \\
&\quad + [C_u \sum_{i=Q+1}^{\infty} (i-Q)p_i - C_u p_{Q+1} - C_u \sum_{i=Q+2}^{\infty} p_i] \\
&= C_o \sum_{i=0}^Q (Q-i)p_i + C_u \sum_{i=Q+1}^{\infty} (i-Q)p_i + C_o \sum_{i=0}^Q p_i - C_u \sum_{i=Q+1}^{\infty} p_i \\
&= E(Q) + C_o F(Q) - C_u (1 - F(Q)) \\
&= E(Q) - C_u + (C_o + C_u) F(Q).
\end{aligned}$$

Therefore we have

$$E(Q) - E(Q + 1) = C_u - (C_o + C_u)F(Q)$$

and

$$E(Q + 1) - E(Q) = -C_u + (C_o + C_u)F(Q).$$

Hence we get (3.7)

$$E(Q) - E(Q - 1) = -C_u + (C_o + C_u)F(Q - 1).$$

We note that

$$\begin{aligned} E(Q) \leq E(Q + 1) &\Leftrightarrow E(Q) - E(Q + 1) \leq 0 \\ &\Leftrightarrow C_u - (C_o + C_u)F(Q) \leq 0 \\ &\Leftrightarrow \frac{C_u}{C_o + C_u} \leq F(Q) \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} E(Q) < E(Q - 1) &\Leftrightarrow E(Q) - E(Q - 1) < 0 \\ &\Leftrightarrow -C_u + (C_o + C_u)F(Q - 1) < 0 \\ &\Leftrightarrow \frac{C_u}{C_o + C_u} > F(Q - 1). \end{aligned} \tag{3.9}$$

Step 3: We are going to analyze the expected cost.

Since $F(Q)$ is increasing in Q , there exists Q^* such that

$$\frac{C_u}{C_o + C_u} \leq F(Q) \quad \text{for } Q \geq Q^*. \quad (3.10)$$

We assume Q^* to be the least integer such that (3.10) is satisfied. This means that

$$\frac{C_u}{C_o + C_u} > F(Q^* - 1).$$

Now we have

$$\frac{C_u}{C_o + C_u} \leq F(Q^*) \leq F(Q^* + 1) \leq F(Q^* + 2) \leq \dots \leq .$$

From (3.8) we have

$$E(Q^*) \leq E(Q^* + 1) \leq E(Q^* + 2) \leq \dots \leq .$$

This implies that the optimal Q should be less than or equal to Q^* .

We also note that

$$\frac{C_u}{C_o + C_u} > F(Q^* - 1) \geq F(Q^* - 2) \geq \dots \geq F(0).$$

From (3.9) we have

$$E(Q^*) < E(Q^* - 1) < E(Q^* - 2) < \dots < E(0).$$

This means that the optimal Q should be greater than or equal to Q^* .

We conclude that the optimal value of Q satisfies

$$F(Q - 1) < \frac{C_u}{C_u + C_o} \leq F(Q).$$

Remark 3.10. In particular if

$$\frac{C_u}{C_o + C_u} \leq F(0)$$

then the optimal ordering quantity Q will be 0.

Example 3.3. Suppose the daily demand follows a Geometric distribution:

$$p_i = (1 - p)p^i, i = 0, 1, 2, \dots \quad \text{where } 0 < p < 1$$

and the overage cost and the underage cost are equal. What is the optimal ordering quantity?

We have

$$F(Q) = \sum_{i=0}^Q p_i = \sum_{i=0}^Q (1 - p)p^i = 1 - p^{Q+1}.$$

Thus the optimal Q should satisfy

$$F(Q - 1) = 1 - p^Q < \frac{1}{2} \leq 1 - p^{Q+1} = F(Q)$$

or equivalently

$$p^{Q+1} \leq \frac{1}{2} < p^Q.$$

Therefore

$$Q^* = \text{integral part of } \frac{\log(0.5)}{\log(p)}.$$

3.6.5 Newsboy Model with Continuous Stochastic Demand

What will be the case when the demand follows a continuous PDF? We expect to get a similar result as in the previous section.

We replace the discrete probability function p_i by a continuous probability density function $f(x)$. Here for simplicity we assume

$$f(x) > 0 \quad \text{for } x \geq 0 \quad \text{and} \quad f(x) = 0 \quad \text{for } x < 0.$$

Proposition 3.14. *Define the cumulative probability function of the demand D to be*

$$F(Q) = \int_0^Q f(t)dt$$

which is strictly increasing in Q . Then the optimal value of Q satisfies

$$F(Q) = \frac{C_u}{C_u + C_o}.$$

Step 1: The expected cost when the ordering quantity is Q :

$$E(Q) = \underbrace{C_o \int_0^Q (Q - x) f(x) dx}_{\text{Expected Overage Cost}} + \underbrace{C_u \int_Q^\infty (x - Q) f(x) dx}_{\text{Expected Shortage Cost}}.$$

We are going to show that $E(Q)$ is minimized when Q is chosen to be the root of the equation

$$F(Q) = \frac{C_u}{C_o + C_u}.$$

Step 2: To find the critical point of the expected cost.

Rewrite

$$\begin{aligned} E(Q) &= C_o Q \int_0^Q f(x) dx - C_o \int_0^Q x f(x) dx \\ &\quad + C_u \int_Q^\infty x f(x) dx - C_u Q \int_Q^\infty f(x) dx. \end{aligned}$$

We note that

$$\frac{d}{dQ} \left\{ Q \int_0^Q f(x) dx \right\} = \int_0^Q f(x) dx + Q f(Q).$$

$$\frac{d}{dQ} \left\{ \int_0^Q x f(x) dx \right\} = Q f(Q).$$

$$\frac{d}{dQ} \left\{ \int_Q^\infty x f(x) dx \right\} = -Q f(Q).$$

$$\frac{d}{dQ} \left\{ Q \int_Q^\infty f(x) dx \right\} = \int_Q^\infty f(x) dx - Q f(Q).$$

and we have

$$\begin{aligned} E'(Q) &= C_o \int_0^Q f(x) dx + C_o Q f(Q) - C_o Q f(Q) \\ &\quad - C_u Q f(Q) - C_u \int_Q^\infty f(x) dx + C_u Q f(Q). \end{aligned}$$

Finally we get

$$\begin{aligned} E'(Q) &= C_o \int_0^Q f(x) dx - C_u \int_Q^\infty f(x) dx \\ &= (C_o + C_u) F(Q) - C_u. \end{aligned}$$

The critical point satisfies the equation

$$F(Q) = \frac{C_u}{C_o + C_u}.$$

Remark 3.11. Since $F(Q)$ is strictly increasing, and $F(0) = 0$ and $F(Q) \rightarrow 1$ as $Q \rightarrow \infty$, the following equation has an unique root

$$F(Q) = \frac{C_u}{C_o + C_u}.$$

Step 3: Analyze the critical point. Since

$$E''(Q) = (C_o + C_u)f(Q) > 0 \quad \text{for } Q > 0,$$

the optimal value of Q satisfies the equation

$$(C_o + C_u)F(Q) - C_u = 0$$

or

$$F(Q) = \frac{C_u}{C_o + C_u}.$$

Example 3.4. If the demand follows the exponential distribution

$$f(x) = \lambda e^{-\lambda x}$$

then the optimal Q will satisfy

$$\begin{aligned} F(Q) &= \int_0^Q \lambda e^{-\lambda x} dx \\ &= 1 - e^{-\lambda Q} \\ &= \frac{C_u}{C_o + C_u}. \end{aligned}$$

Thus the optimal Q is

$$\frac{-1}{\lambda} \log \frac{C_o}{C_o + C_u}.$$

Remark 3.12. We give some remarks on the newsboy model.

1. It is clear that the optimal ordering quantity increases (decreases) as the underage cost C_u increases (overage cost C_o increases).
2. In the discrete case, it is difficult to determine the optimal Q when the demand is a Poisson process with probability distribution

$$\frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

- But when λ is large, the Poisson distribution can be approximated by the normal distribution with mean and variance λ .
- So by using the result in continuous case, one can get an approximated optimal ordering quantity very quickly when λ is large.

3.7 A Summary of Learning Outcomes

- Able to **give** the definitions of a Poisson Process and the Markov property.
- Able to **state and prove** the relationships among a Poisson process, the exponential distribution and the Poisson distribution.
- Able to **state and prove** the relationships among the Markov property, the exponential distribution and the Erlangian distribution.
- Able to **formulate** Newsboy's problem and **solve** for its optimal policy.

3.8 Exercises

1. Consider the geometric distribution

$$P(x = i) = p_0(1 - p_0)^i, \quad 0 < p_0 < 1 \quad i = 0, 1, \dots, \infty$$

Prove that for any two non-negative integers t and h , we have

$$P(x \geq t + h | x \geq t) = P(x \geq h).$$

2. Consider an non-negative discrete random variable x taking values in $\{0, 1, 2, \dots\}$ such that $P(x = i) = p_i, i = 0, 1, 2, \dots$ and $0 < p_0 < 1$. Suppose that for any two non-negative integers t and h , we have

$$P(x \geq t + h | x \geq t) = P(x \geq h)$$

prove that

$$p_i = p_0(1 - p_0)^i, \quad i = 0, 1, \dots, \infty$$

3. Let $p(h) = e^{-\lambda h} - 1 + \lambda h$, prove that

$$\lim_{h \rightarrow 0} \frac{p(h)}{h} = 0$$

and hence deduce that $e^{-\lambda h} = 1 - \lambda h + o(h)$.

4. Events occur according to a Poisson process with rate $\lambda = 2$ per hour.
- (a) What is the probability that no event occurs between 8 p.m. and 9 p.m.?
 - (b) Starting at noon, what is the expected time at which the fourth event occurs?
 - (c) Find the probability that two or more events occur between 6 p.m. and 8 p.m..
5. Customers arrive at a bank at a Poisson rate λ . Suppose two customers arrived during the first hour. What is the probability that
- (a) both arrived during the first 20 minutes?
 - (b) at least one arrived during the first 20 minutes?

3.9 Suggested Solutions

1. We note that

$$P(x \geq h + t) = \sum_{i=t+h}^{\infty} p_0(1 - p_0)^i = (1 - p_0)^{h+t},$$

$$P(x \geq t) = \sum_{i=t}^{\infty} p_0(1 - p_0)^i = (1 - p_0)^t,$$

and

$$P(x \geq h) = \sum_{i=h}^{\infty} p_0(1 - p_0)^i = (1 - p_0)^h.$$

Hence

$$\frac{P(x \geq t + h)}{P(x \geq t)} = \frac{(1 - p_0)^{h+t}}{(1 - p_0)^t} = (1 - p_0)^h = P(x \geq h).$$

The result follows.

2. Now from the given condition we have

$$\frac{\sum_{i=t+h}^{\infty} p_i}{\sum_{i=t}^{\infty} p_i} = \sum_{i=h}^{\infty} p_i.$$

Put $t = 1$ we have

$$\sum_{i=h+1}^{\infty} p_i = (1 - p_0) \sum_{i=h}^{\infty} p_i$$

and therefore

$$p_h = p_0 \sum_{i=h}^{\infty} p_i.$$

Now put $h = 1$ we have

$$p_1 = p_0(1 - p_0).$$

Put $h = 2$ we have

$$p_2 = p_0(1 - p_0 - p_1) = p_0(1 - p_0 - p_0(1 - p_0)) = p_0(1 - p_0)^2.$$

Inductively (or by using Mathematical Induction), the result follows.

3. By using L'hospital rule we have

$$\lim_{h \rightarrow 0} \frac{e^{-\lambda h} - 1 + \lambda h}{h} = \lim_{h \rightarrow 0} \frac{-\lambda e^{-\lambda h} + \lambda}{1} = 0.$$

Therefore by definition we have $p(h) = o(h)$. Hence

$$o(h) = e^{-\lambda h} - 1 + \lambda h$$

and the result follows.

4. Let $X(t)$ be the number of occurrences in t hours.

(a) $P(X(1) = 0) = e^{-2}$

(b) Let T_n denote the elapsed time between the $(n - 1)$ th and n th event.

$$E(T_1 + T_2 + T_3 + T_4) = 4\left(\frac{1}{\lambda}\right) = 2$$

Therefore, the expected time at which the fourth event occurs is 2 P.M.

(c) $P(\text{Two or more events occur between 6 P.M. and 8 P.M.})$

$$= 1 - P(X(2) = 0) - P(X(2) = 1)$$

$$= 1 - e^{-4} - 4e^{-4}$$

$$= 1 - 5e^{-5}$$