

# Discrete Probability Distribution

---

PART I

# Learning Outcomes

---

## Learning Outcomes

At the end of this lesson, you will be able to:

1. Identify characteristics of a probability distribution.
2. Calculate the mean, variance and expected value of a discrete probability distribution.
3. Identify characteristics of a Binomial and Poisson distribution through the assumptions for each distribution.
4. Calculate probabilities, mean and standard deviation of the Binomial and Poisson distribution using relevant formulas.
5. Solve real-life business problems by applying concepts of discrete, Binomial and Poisson distribution.

# Introduction to Discrete Probability Distribution

- Your company's network keeps getting disruptions on a particular day. Based on past incidents, you have collected data on number of network disruptions and possible outcomes:

Outcomes:	Disruptions/day	Probability
0		0.35
1		0.25

- Your boss wants to know what is the average network disruptions per day. How would you find out?



- This case study is an example of a discrete probability distribution.
- In this topic, you will apply concepts of discrete probability distribution to solve real life problems.

# What is a Discrete Probability Distribution?

---

- A Discrete Random Variable is one that takes on countable values.
- A **Discrete Probability Distribution** is a table that shows each value of random variable and its probability of occurrence.
- Characteristics of a Discrete Probability Distribution
  - ✓ The probability of a particular outcome is between 0 and 1 inclusive.
  - ✓ The outcomes are mutually exclusive.
  - ✓ Sum of the probabilities of the outcomes is equal to 1.

# What is a Probability Distribution?

Example: The Probability Distribution of  $x$ , number of network interruptions per day is shown below:

$x=k$	0	1	2	3	4	5
$P(x=k)$	0.35	0.25	0.20	0.10	0.05	0.05

where

- $x$ = number of network interruptions per day
- $P(x=k)$  is the probability of outcome

Notice that:

- Probability of a particular outcome is between 0 and 1 inclusive.
- Sum of the probabilities of the outcomes is equal to 1.

# What is a Probability Distribution?

- Example 1: Explain whether the following is a discrete probability distribution function.

X=k	1	2	3	4
P(X=k)	0.09	0.36	0.49	0.05

$$0 \leq P(X=k) \leq 1$$

$$\sum_{k=1}^4 P(X = k) = 0.09 + 0.36 + 0.49 + 0.05 = 0.99 \neq 1$$

Since sum of probability is not 1, it is NOT a discrete probability distribution function.

# Mean and Variance of a Probability Distribution

---

## Mean, $\mu$

- A typical value used to represent the central location of the data.
- Also referred to as the expected value

Mean of a Probability Distribution,  $\mu = \sum [k \cdot P(X = k)]$

## Variance, $\sigma^2$

- Amount of spread (or variation)

Variance of a Probability Distribution,  $\sigma^2 = [\sum k^2 \cdot P(X = k)] - \mu^2$

## Standard deviation

- Positive square root of the variance  $= \sqrt{\sigma^2} = \sigma$

# Mean and Variance of a Probability Distribution

---

Example 2: Find the mean, variance and standard deviation of the random variable in the following probability distribution:

$X = k$	1	2	3	4	5
$P(X = k)$	0.16	0.22	0.28	0.20	0.14

$$\text{Mean, } \mu = \sum kP(X=k) = 1(0.16) + 2(0.22) + 3(0.28) + 4(0.20) + 5(0.14) = 2.94$$

$$\sum k^2 P(X=k) = 1^2(0.16) + 2^2(0.22) + 3^2(0.28) + 4^2(0.20) + 5^2(0.14) = 10.26$$

$$\text{Variance, } \sigma^2 = \sum k^2 P(X=k) - \mu^2 = 10.26 - 2.94^2 = 1.6164$$

$$\text{Standard deviation, } \sigma = \sqrt{1.6164} = 1.27$$

# Types of Discrete Probability Distribution

---

- There are a variety of discrete probability distributions that you can use to model different types of data. The correct discrete distribution depends on the properties of your data. For example, use
  - ✓ Binomial distribution to model binary data, such as coin tosses.
  - ✓ Poisson distribution to model count data, such as the count of library book checkouts per hour.

---

PART II

# Binomial Distribution

---

# Introduction to Binomial Distribution

- You work for an online store selling shoes.
- Recent statistics suggest that 15% of those who visit your company website will make a purchase.
- A sample of 16 ‘hits’ were selected, and it was found that 4 purchases were made.
- What is the probability of exactly 4 purchases are made?



- The above case study is an example of Binomial Distribution.
- One of its characteristics is that there are only 2 outcomes :  
Probability (purchase) = 0.15  
Probability (No purchase made) =  $1 - 0.15 = 0.85$
- In this topic, you will apply concepts of Binomial Distribution to solve real life problems.

# Binomial Distribution

- Binomial distribution is a common type of discrete probability distribution.
- What type of random experiment qualifies to be a binomial experiment?

A binomial experiment whose intent is to know ‘the number of trials which is successful’ has to satisfy the following conditions:

- ✓ A trial is repeated for a fixed number of times in the experiment.
- ✓ Each trial has exactly 2 outcomes: success or failure.
- ✓ The probability of success (or failure) is the same for each trial.
- ✓ The trials are independent.

# Binomial Distribution

Let  $X$  be the random variable ‘the number of successes in  $n$  trials’ in a binomial experiment  $X \sim Bin(n, p)$

The probability of exactly  $x$  successes in  $n$  trials is given by:

$$P(X = x) = nCx p^x (1 - p)^{n-x} \text{ where } x = 0, 1, 2, \dots, n$$

$n$  : number of times a trial is repeated

$p$  : probability of success in a single trial

$x$  : the random variable that represents the number of successes in  $n$  trials.

# Binomial Distribution

Example 3: Recent statistics suggest that 15% of those who visit your company website will make a purchase. A sample of 16 'hits' were selected, and it was found that 4 purchases were made.

What is the probability of exactly 4 purchases are made?

$$\begin{aligned}X &\sim B(n, p) \\X &\sim B(16, 0.15)\end{aligned}$$

n : number of times a trial is repeated  
p : probability of success in a single trial  
x : the random variable that represents the number of successes in n trials.

$$\begin{aligned}n &= 16, \quad p = 0.15, \quad x = 4 \\P(x) &= {}_n C_x p^x (1 - p)^{n-x} \\P(4) &= {}_{16} C_4 0.15^4 (1 - 0.15)^{16-4} = 0.1311\end{aligned}$$

$$P(\text{Exactly 4 purchases made}) = 0.1311$$

## Binomial Distribution

---

Example 4: Microfracture knee surgery has a 75 % chance of success on patients with degenerative knees. The surgery is performed on three patients. Find the probability of the surgery being successful on exactly two patients.

$$\begin{aligned}X &\sim B(n, p) \\X &\sim B(3, 0.75)\end{aligned}$$

n : number of times a trial is repeated  
p : probability of success in a single trial  
x : the random variable that represents the number of successes in n trials.

$$\begin{aligned}n &= 3, \quad p = 0.75, \quad x = 2 \\P(x) &= {}_n C_x p^x (1-p)^{n-x} \\P(2) &= {}_3 C_2 0.75^2 (1 - 0.75)^{3-2} = 0.422\end{aligned}$$

$$P(\text{Successful on exactly two patients}) = 0.422$$

# Mean and Variance of a Binomial Distribution

For Binomial Distribution,

Expectation or population mean,  $\mu$  is given by

$$\mu = np$$

Population variance,  $\sigma^2$  is given by

$$\sigma^2 = np(1-p)$$

$n$  : number of times a trial is repeated

$p$  : probability of success in a single trial

$q$  : probability of failure in a single trial

# Mean and Variance of a Binomial Distribution

Example 5: 5% of workers at construction sites are known to suffer from hearing impaired problem due to the unhealthy noise level. If we randomly select 28 workers from construction sites, find

- a) probability that exactly 4 of them suffer from hearing impaired problem
- b) mean and standard deviation of the number of workers suffering from hearing impaired problem.

$$\begin{aligned}X &\sim B(n, p) \\X &\sim B(28, 0.05)\end{aligned}$$

$$\begin{aligned}a) \quad P(k) &= {}_n C_k p^k (1-p)^{n-k} \\P(X = 4) &= {}_{28} C_4 0.05^4 (1-0.05)^{28-4} \\&= (20,475) 0.05^4 (0.95)^{24} = 0.0374\end{aligned}$$

$$P(\text{Exactly 4 has hearing impaired problem}) = 0.0374$$

## Mean and Variance of a Binomial Distribution

---

$$\begin{aligned} b) \quad \mu &= np = (28)(0.05) = 1.4 \\ \sigma^2 &= np(1 - p) = (1.4)(0.95) = 1.33 \\ \sigma &= 1.1532 \end{aligned}$$

# Poisson Distribution

---

PART III

# Introduction to Poisson Distribution

- You are working in the Traffic Police Department.
- Based on known data, you know the average number of accidents per month at traffic junction at Sin Min road is 3.
- What is the probability in any given month that 4 accidents will happen at that junction?



- The above case study is an example of Poisson Distribution.
- A Poisson distribution is a statistical distribution showing the likely number of times that an event will occur within a specified period of time.
- In this topic, you will apply concepts of Poisson Distribution to solve real life problems.

# Poisson Distribution

A Poisson experiment whose intent is to know ‘the number of occurrences in an interval’ has to satisfy the following conditions:

- ✓ The experiment counts the number of times an event occurs in a given interval.
- ✓ The interval can be an interval of time, area, or volume.
- ✓ The probability of the event occurring is the same for each interval.
- ✓ The number of occurrences in one interval is independent of the number of occurrences in other intervals.

# Poisson Distribution

Let  $X$  be the random variable ‘the number of occurrences in an interval’.

$$X \sim Po(\mu)$$

The probability of exactly  $x$  occurrences in an interval is given by:

$$P(X = k) = e^{-\mu} \frac{\mu^k}{k!} \text{ where } k = 0, 1, 2, \dots$$

Where  $X$  = number of events occurring per unit interval

$\mu$  = mean rate of occurrences per unit interval

$e$  = Euler’s number,  $e$  approx. equal to **2.718**

# Poisson Distribution

Example 6: Mean number of accidents per month at traffic junction X is 3. What is the probability in any given month

- a) 4 accidents will happen at the junction
- b) more than 2 accidents will happen at the junction

where

$X$  = number of events occurring in an interval

$\mu$  = mean rate

$e$  = constant equal to 2.718

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$
$$\text{a) } P(4) = \frac{3^4 e^{-3}}{4!} = 0.1680$$

$P(4 \text{ accidents will happen at the junction}) = 0.1680$

# Poisson Distribution

---

$$\begin{aligned} b) P(x > 2) &= 1 - (P(x = 0) + P(x = 1) + P(x = 2)) \\ &= 1 - \left( \frac{3^0 e^{-3}}{0!} + \frac{3^1 e^{-3}}{1!} + \frac{3^2 e^{-3}}{2!} \right) \\ &= 1 - (0.0498 + 0.1494 + 0.2240) = 0.5768 \end{aligned}$$

$P(\text{More than 2 accidents will happen at the junction}) = 0.5768$

# Poisson Distribution

---

Example 7: A car salesman sells on the average 3 cars per week (assume 5 working days per week). What is the probability that in a given week, he will sell some cars?

$$\begin{aligned}P(x) &= \frac{\mu^x e^{-\mu}}{x!} \\P(x > 0) &= 1 - P(x = 0) \\&= 1 - \frac{3^0 e^{-3}}{0!} = 0.9502\end{aligned}$$

$$P(\text{Salesman will sell some cars}) = 0.9502$$

# Mean and Variance of Poisson Distribution

---

For a Poisson random variable,

- ✓ Expectation or population mean =  $\mu$
- ✓ Population Variance,  $\sigma^2 = \mu$

# Mean and Variance of Poisson Distribution

---

Example 8: On average, 2 accidents take place every hour in a factory.

- a) What is the probability that at most 2 accidents take place in three hours.
- b) Compute the expected number of accidents and the standard deviation in 4 hours

a) Average number of accidents per hour,  $\mu = 2$

Average number of accidents per 3 hours = 6

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

$$\begin{aligned}P(x \leq 2) &= P(x = 0) + P(x = 1) + P(x = 2) \\&= \frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!} + \frac{6^2 e^{-6}}{2!} = 0.0620\end{aligned}$$

P(at most 2 accidents take place in three hours) = 0.0620

# Mean and Variance of Poisson Distribution

---

b) Expected number of accidents = 8 hours

$$\sigma = \sqrt{8} = 2.8248$$