# Poisson Distribution

# Poisson Distribution

**Why do we need Poisson distribution?**

In Binomial distribution, we use the value of n (independent experiments) and p (probability of events). But if n is large or tends to $\infty$ and p is so small or tends to 0. Then it is very cumbersome to know the probability distribution by any other method or by Binomial distribution.

So, we need to use Poisson distribution.

Poisson distribution is, in fact, a limiting case of Binomial distribution.

# Poisson Distribution

Suppose that the average rate of transcription is one mRNA per second. Then,

**What is the probability that eight mRNAs will be produced in 10 seconds?**

In deterministic systems, we should get 10 mRNAs in 10 seconds. However, since the transcription process is stochastic in nature, i.e. transcription (mRNA) may happen or not. Thus, we need to calculate probability of having eight mRNAs in 10 seconds.

Let us try to use the Binomial distribution to calculate the probability.

For that, let us take the discrete time,

$$\Delta t = 1 \text{ second.}$$

It means that that there are 10 time points in 10 second, i.e. $N = 10$.

It means that at each trial, either mRNA is produced or nothing happens.

# Poisson Distribution

For Binomial distribution, the average number of mRNA produced

$$\mu = N \times p = 10$$

$$\Rightarrow p = 1.$$

Thus,

$$P(X=k) = {}^{N}C_{k}\,(p)^{k}\,(1-p)^{N-k}$$

In this case, k = 8, then

$$P(X=8) = {}^{10}C_{8}\,(p)^{8}\,(1-p)^{10-8} = 0 \text{ (as } 1-p = 0).$$

But, this is unrealistic value.

# Poisson Distribution

Let us try again using Binomial distribution, but with different time intervals.

Now, let us consider, $\Delta t = 0.1$ second

Then the number of trials in 10 seconds, $N = 100$

The average number of mRNA produced,

$$\mu = N \times p = 10$$
$$\Rightarrow p = 10/100 = 0.1$$

It implies that,

$$P(X=k) = {}^{N}C_{k} \, (p)^{k} \, (1-p)^{N-k}$$

Here, k = 8, then

$$P(X=8) = {}^{100}C_{8} \, (0.1)^{8} \, (1-0.1)^{92}$$
$$\Rightarrow P(X=8) = 0.1148$$

Although this value is realistic, it changes with different time intervals.

# Poisson Distribution

Let us try again using Binomial distribution, but with different time intervals.

Now, let us consider, $\Delta t = 0.01$ second

Then the number of trials in 10 seconds, $N = 1000$

The average number of mRNA produced,

$$\mu = N \times p = 10$$
$$\Rightarrow p = 10/1000 = 0.01$$

It implies that,

$$P(X=k) = {}^{N}C_{k}\, (p)^{k}\, (1-p)^{N-k}$$

Here, k = 8, then

$$P(X=8) = {}^{1000}C_{8}\, (0.01)^{8}\, (1-0.01)^{992}$$
$$\Rightarrow P(X=8) = 0.1126xx$$

Although this value is realistic, it changes with different time intervals.

# Poisson Distribution

Let us try again using Binomial distribution, but with different time intervals.

Now, let us consider, $\Delta t = 0.001$ second (i.e. 1 milli second)

Then the number of trials in 10 seconds, $N = 10000$

The average number of mRNA produced,

$$\mu = N \times p = 10$$
$$\Rightarrow p = 10/10000 = 0.001$$

It implies that,

$$P(X=k) = {}^{N}C_{k}\,(p)^{k}\,(1-p)^{N-k}$$

Here, $k = 8$, then

$$P(X=8) = {}^{10000}C_{8}\,(0.001)^{8}\,(1-0.001)^{9992}$$
$$\Rightarrow P(X=8) = 0.1126xx$$

Although this value is realistic, it changes with different time intervals.

# Poisson Distribution

We see that the value tends to a realist constant value with decrease in the time interval.

So, actually the problem is that the time is continuous, thus discretization of time leads to error.

The solution to this problem is to make the time interval very small, i.e. Δt -> 0.

That is, N tends to infinity and p values tends to zero (i.e. very small).

We can re-write Binomial distribution as

$$P(X = k) = \lim_{N \to \infty} {}^{N}C_k (p)^k (1 - p)^{N-k}$$

Substituting p as (μ/N), we get

$$P(X = k) = \lim_{N \to \infty} {}^{N}C_k \left(\frac{\mu}{N}\right)^k \left(1 - \frac{\mu}{N}\right)^{N-k}$$

# Poisson Distribution

Re-arranging the factors gives the following equation,

$$P(X=k)=\lim_{N\to\infty}\frac{N(N-1)\cdots(N-k+1)}{k!}\left(\frac{\mu}{N}\right)^k\left(1-\frac{\mu}{N}\right)^N\left(1-\frac{\mu}{N}\right)^{-k}$$

Taking out N-independent factors outside gives the following equation,

$$P(X=k)=\frac{\mu^k}{k!}\lim_{N\to\infty}\frac{N(N-1)\cdots(N-k+1)}{N^k}\left(1-\frac{\mu}{k}\right)^N\left(1-\frac{\mu}{N}\right)^{-k}$$

Re-arranging the factors further gives the following equation,

$$p(X=k)=\frac{\mu^k}{k!}\times\frac{N^k}{N^k}\lim_{N\to\infty}\left(1-\frac{\mu}{N}\right)^N\lim_{N\to\infty}\left(1-\frac{\mu}{N}\right)^{-k}$$

# Poisson Distribution

By substituting the value of limits, we get

$$P(X = k) = \frac{\mu^k}{k!} \times 1 \times \mathrm{e}^{-\mu} \times 1$$

It implies that

$$P(X = k) = \frac{\mu^k}{k!} \mathrm{e}^{-\mu}$$

This equation is independent of N and Δt. This is known as **Poisson distribution**.

Using the Poisson distribution, we get the probability of having 8 mRNAs

$$P(X = 8) = \frac{(10)^8}{8!} \mathrm{e}^{-10} = 0.1126$$

# Poisson Distribution

It is the discrete probability distribution of the number of events occurring in a given time period, given the average number of times the event occurs over that time period.

It is used to show how many times an event occurs over a specific period.

It is the distribution related to probabilities of events that are extremely rare but have a large number of independent opportunities for occurrence.

**The key characteristics of a Poisson distribution are as follows:**

**Discreteness:** The Poisson distribution deals with **discrete random variables,** meaning the number of events can only take on non-negative integer values (0, 1, 2, 3,…).

**Independence:** Each event in Poisson distribution is considered to be independent of others, meaning the occurrence of any event doesn't affect the occurrence of another event in the same interval.

**Constant rate:** The events are assumed to occur at a constant average rate ($\mu$) over the given interval.

# Poisson Distribution

Let us take an example of DNA replication. Consider that each time DNA is replicated, the rate (frequency) of errors in replication is μ.

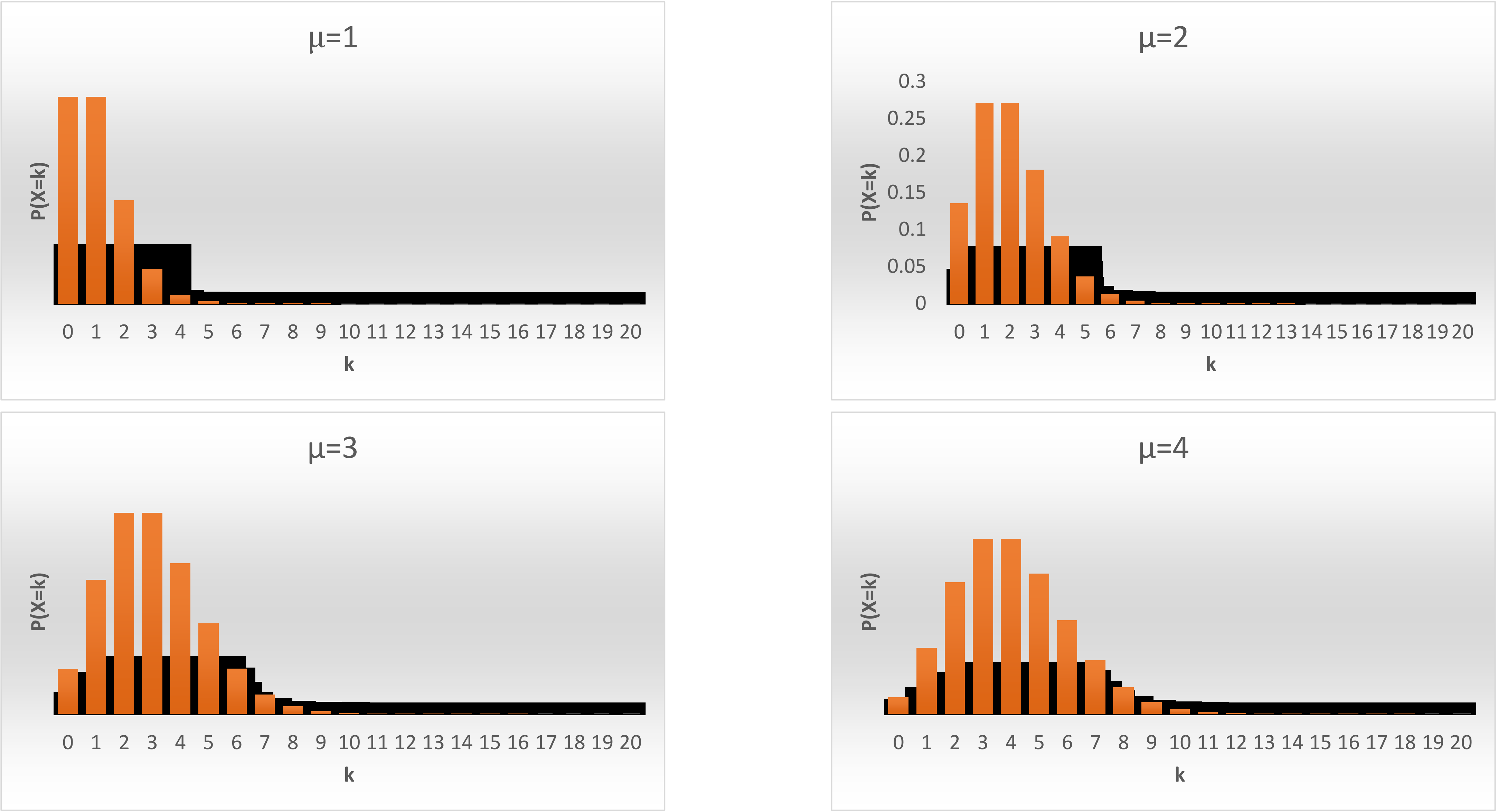
Also, let us assume that the length of the DNA is L (e.g. 2.9 Gbp or 2.9 x $10^9$ bp).

What is the probability of having k errors in the genome after replication?

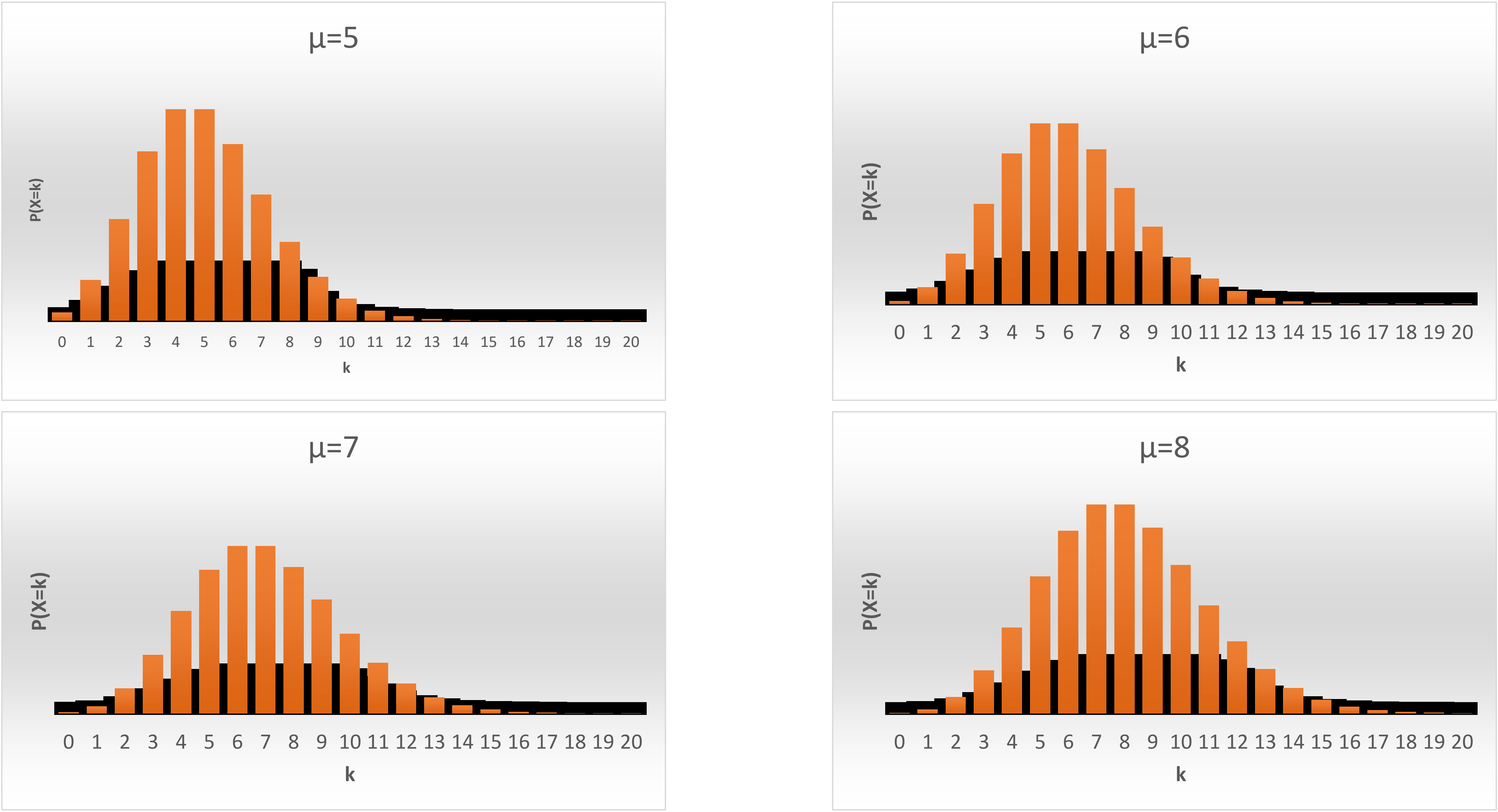
$$P(\mathrm{X} = k) = \frac{\mu^k}{k!} \mathrm{e}^{-\mu}$$

where, **P(X = k)** is the probability that an event will occur k times, **X** is a random variable following a Poisson distribution, **μ** is the average number of times an event occurs, k is the number of times an event occurs.

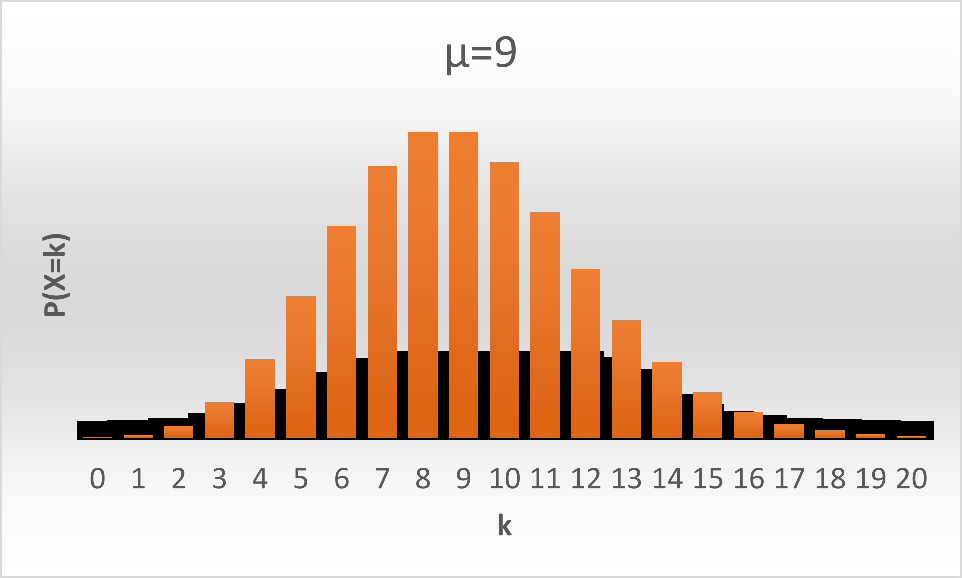
# Poisson Distribution

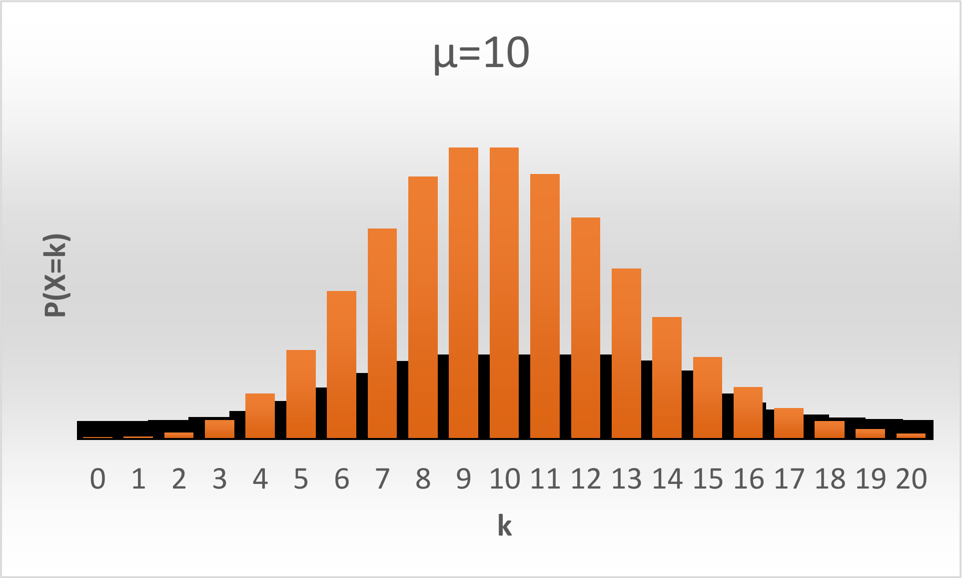


# Poisson Distribution



# Poisson Distribution



**Poisson Distribution**

# Poisson Distribution

## Mean of Poisson Distribution

$$E(X) = \mu = np$$

where, **E[X]** is the mean of the Poisson's Distribution, **μ** is the parameter of the distribution, **X** is a random variable following a Poisson distribution, **n** is the number of trails, and **p** is the probability of success.

## Variance of Poisson Distribution

$$Var(X) = \mu$$

# Poisson Distribution

## Difference between Binomial and Poisson Distribution

| Aspect | Binomial Distribution | Poisson Distribution |
|---|---|---|
| Nature | Discrete | Discrete |
| Number of Trials | Fixed (n) | Unlimited |
| Outcome | Success or Failure | Rare Events |
| Parameter | Probability of Success (p) | Average Event Rate ($\lambda$) |
| Possible Values | 0 to n | 0, 1, 2, . . . |
| Mean | $\mu = n \times p$ | $\mu = \lambda$ |
| Variance | $\sigma^2 = n \times p \times (1 - p)$ | $\sigma^2 = \lambda$ |
| Applicability | Limited to a fixed number of trials | Rare events over a large population |
| Example | Flipping a coin multiple times | Counting occurrences of an event |
| Assumptions | Independent trials, constant p | Rare events, low probability of success |

# Poisson Distribution

**Example 1:**
**Suppose there is a bakery on the corner of the street and on average 10 customers arrive at the bakery per hour. For this case, we can calculate the probabilities of different numbers of customers arriving at the bakery at any hour using the Poisson distribution.**

Probability of having exactly 5 customers arrive in an hour:
$P(X = 5) = 10^5 \times e^{-10}/5! \approx 0.037$

Probability of having no customers arrive in an hour:
$P(X = 0) = 10^0 \times e^{-10}/0! \approx 4.54 \times 10^{-5}$

Probability of having at least 15 customers arrive in an hour (sum of probabilities from 15 to infinity):
$P(X \geq 15) = 1 – P(X < 15) = 1 – (P(X = 0) + P(X = 1) + . . . + P(X = 14))$

# Poisson Distribution

**Example 2:**
**If 4% of the total items made by a factory are defective. Find the probability that less than 2 items are defective in the sample of 50 items.**

Here we have, $n = 50$, $p = (4/100) = 0.04$, => $\mu = np = 2$

Using Poisson's Distribution,

$P(X = 0) = \frac{2^0 e^{-2}}{0!} = 1/e^2 = 0.13534$

$P(X = 1) = \frac{2^1 e^{-2}}{1!} = 2/e^2 = 0.27068$

Hence the probability that less than 2 items are defective in sample of 50 items is given by:

$P( X < 2 ) = P( X = 0 ) + P( X = 1 ) = 0.13534 + 0.27068 = 0.40602$

# Poisson Distribution

**Example 3:**

**If the probability of a bad reaction from medicine is 0.002, determine the chance that out of 1000 persons more than 3 will suffer a bad reaction from medicine.**

Here we have, $n = 1000$, $p = 0.002$, $\mu = np = 2$

X = Number of person suffer a bad reaction

Using Poisson's Distribution

$P(X > 3) = 1 - \{P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)\}$

$P(X = 0) = \frac{2^0 e^{-2}}{0!} = 1/e^2$

$P(X = 1) = \frac{2^1 e^{-2}}{1!} = 2/e^2$

$P(X = 2) = \frac{2^2 e^{-2}}{2!} = 2/e^2$

$P(X = 3) = \frac{2^3 e^{-2}}{3!} = 4/3e^2$

$P(X > 3) = 1 - [19/3e^2] = 1 - 0.85712 = 0.1428$

# Poisson Distribution

**Example 4:**
**If 1% of the total screws made by a factory are defective. Find the probability that less than 3 screws are defective in a sample of 100 screws.**

Here we have, n = 100, p = 0.01 => $\mu$ = np = 1
X = Number of defective screws

Using Poisson's Distribution
P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)
P(X = 0) = $\frac{1^0 e^{-1}}{0!}$ = 1/e
P(X = 1) = $\frac{1^1 e^{-1}}{1!}$ =1/e
P(X = 2) = $\frac{1^2 e^{-1}}{2!}$ =1/2e

P(X < 3) = 1/e + 1/e +1/2e
= 2.5/e = 0.919698

# Poisson Distribution

**Example 5:**
**If in an industry there is a chance that 5% of the employees will suffer by corona. What is the probability that in a group of 20 employees, more than 3 employees will suffer from the corona?**

Here we have, $n = 20$, $p = 0.05$, => $\mu = np = 1$
X = Number of employees who will suffer corona

Using Poisson's Distribution
$P(X > 3) = 1-[P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]$

$P(X = 0) = \frac{1^0 e^{-1}}{0!} = 1/e$

$P(X = 1) = \frac{1^1 e^{-1}}{1!} = 1/e$

$P(X = 2) = \frac{1^2 e^{-1}}{2!} =1/2e$

$P(X = 3) = \frac{1^3 e^{-1}}{3!} =1/6e$

$P(X > 3) = 1 – [1/e + 1/e + 1/2e + 1/6e]$
$= 1 – [ 8/3e] = 0.018988$

# Poisson Distribution

**Example 6:**
**A manufacturer knows that the bulb he makes consists of 2% of bulbs as defective. If he makes 200 bulbs then what is the probability that less than 4 bulbs are defective?**

Here we have, n = 200, p = 0.02 => μ = np = 4
X = Number of bulbs are defective

Using Poisson's Distribution: P(X < 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)

$P(X = 0) = \frac{4^0 e^{-4}}{0!} = 1/e^4$

$P(X = 1) = \frac{4^1 e^{-4}}{1!} = 4/e^4$

$P(X = 2) = \frac{4^2 e^{-4}}{2!} = 8/e^4$

$P(X = 3) = \frac{4^3 e^{-4}}{3!} = 32/3e^4$

$P(X < 4) = 1/e^4 + 4/e^4 + 8/e^4 + 32/3e^4 = 71/3e^4 = 0.43347$

# Poisson Distribution

**Example 7:**
**At a university, the probability that a member of staff is absent on any day is 0.001. If there are 800 members of staff, calculate the probabilities that the number absent on any day is 4.**

Here we have, n = 800, p = 0.001 => μ = np = 0.8
X = Number of members of staff absent on any day

Using Poisson's Distribution

$$P(X = 4) = \frac{(0.8)^4 e^{-0.8}}{4!} = 0.00767$$

# Thank You