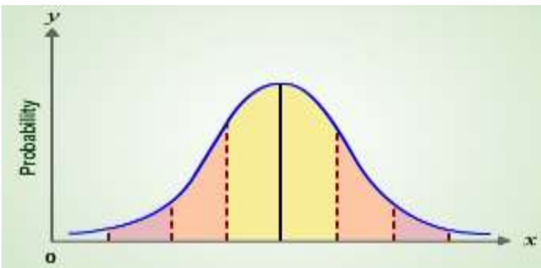
Learning Outcome

By the end of this lecture, you will be able to:

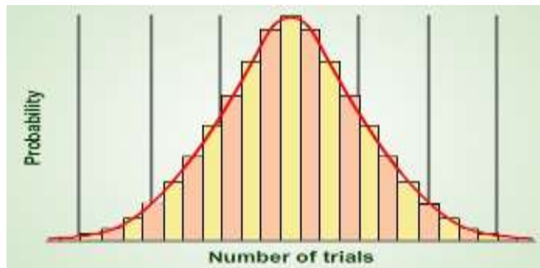
- Define the Hypergeometric and Poisson distributions
- Calculate the probabilities of Hypergeometric and Poisson random variables

Introduction

Recall normal distribution and binomial distribution.



Normal distribution



Binomial distribution

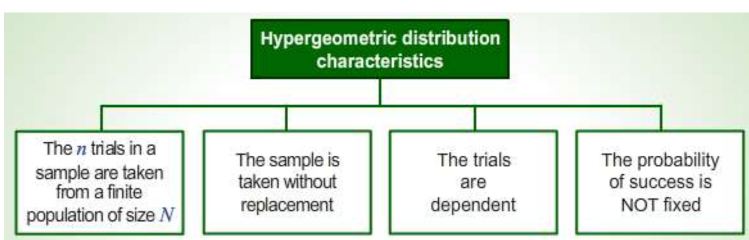
Other than these distributions, we have to use several other distributions depending on nature of sample and experiment. They include:

- Hypergeometric distributions
- Poisson distributions

In this lecture, we will explore Hypergeometric and Poisson distributions and the probabilities of Hypergeometric and Poisson random variables.

Hypergeometric Distribution – Introduction

We'll begin with hypergeometric distribution. Here are the characteristics of the hypergeometric distribution.



Hypergeometric distribution characteristics

If the random variable X counts the number (x) of successes from (K) successes in the population in a sample of size (n) drawn from a population of size (N), then X is said to have the hypergeometric probability distribution.

Hypergeometric Distribution Formula

Given here is the Hypergeometric distribution formula:

$$P(x) = \frac{C_{n-x}^{N-K} \; C_x^K}{C_n^N} \qquad (1)$$

Where:

- N = Population size
- K = Number of successes in the population
- $N - K$ = Number of failures in the population
- n = Sample size
- x = Number of successes in the sample
- $n - x$ = Number of failures in the sample

Example

A sample of three light bulbs was selected, without replacement, from a box of ten. Of the ten, 4 were defectives. Let X denote the number of defective bulbs in a sample of three, and find:

1. The probability that two of the three selected were defective.

Solution:

The probability that two of the three selected were defective is:
P(any two are defective) = $P(X = 2)$ ➔

$$P(X = 2) = \frac{C_{3-2}^{10-4} \; C_2^4}{C_3^{10}}$$

$$= \frac{C_1^6 \; C_2^4}{C_3^{10}}$$

$$= \frac{(6)(6)}{(120)} = 0.3$$

2. The probability that ONLY the second is good.

Solution:

The probability that ONLY the second is good is:

$$P(DGD) = \frac{4}{10}\frac{6}{9}\frac{3}{8}$$
$$= \frac{72}{720} = 0.1$$

3. Write the probability distribution, and identify it, for the r.v. X

The probability distribution for the *r.v.* X:
X has the hypergeometric distribution with $N = 10$, $K = 4$, $n = 3$

$$P(X = x) = \frac{C_{3-x}^{10-4}\ C_x^4}{C_3^{10}}$$

Poisson Probability Distribution

Characteristics of the Poisson Distribution:

1. Random events occur in time or space such that the average number of outcomes of interest per *time* or some *space* is λ.
2. The mean rate of occurrence is constant, λ.
3. Events in disjoint intervals are independent.
4. No two events can occur in a very small interval of time or space

The Poisson probability distribution is given by…

$$P(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}\ ,\quad x = 0,1,2,\ldots \qquad (2)$$

Where:
- x: Number of events in a segment (interval) of a space
- t: Size of the segment of interest (number of segments)
- λ: Expected (average) number of events in a segment of unit size
- e: Base of the natural logarithmic function

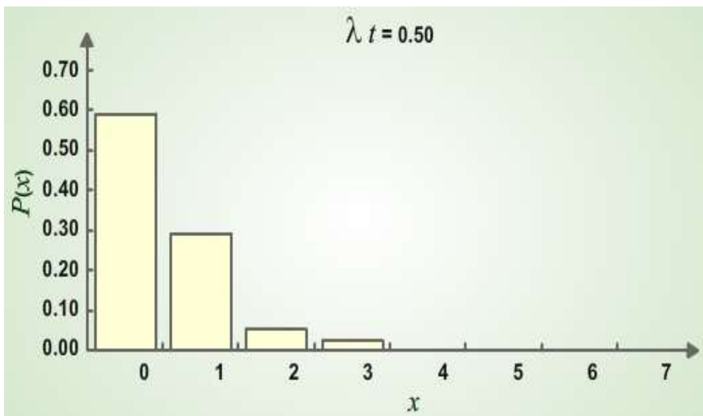
Mean and Variance of Poisson Distribution

Mean and variance of Poisson distribution are calculated by using the formulas given here:

$$Mean \ \ \mu = \lambda t \qquad (3)$$
$$Variance: \sigma^2 = \lambda t \text{ and standard deviation } \ \sigma = \sqrt{\lambda t} \qquad (4)$$

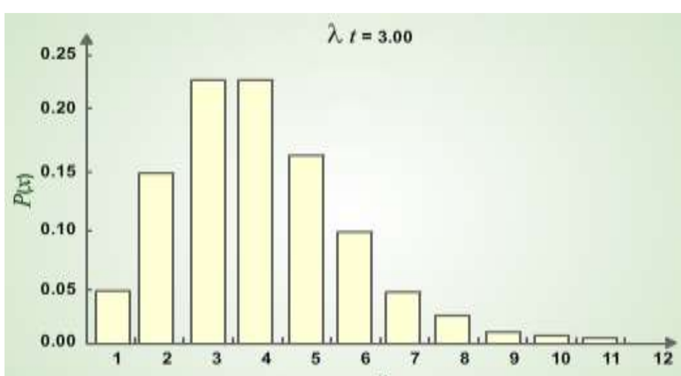
The shape of the Poisson distribution depends on the two parameters λ and t, i.e. the *mode* will always equal to λt:

$\boxed{\lambda t = 0.50}$



Poisson distribution

$\boxed{\lambda t = 3.0}$



Poisson distribution

Example

At a checkout counter, customers arrive at an average of 2.5 customers per minute. If *X* denotes the number of customers arriving at the counter per minute, find:

1. The probability distributions of the *r.v. X* and identify it.

Solution:
X has the Poisson distribution with λ=2.5 *and t*=1 min. The probability function is given by

$$P(X=x)=\frac{2.5^x e^{-2.5}}{x!}\ ,\quad x=0,1,2,\ldots$$

2. The probability that two will arrive within 2 minutes.

Solution:
The probability that two will arrive within 2 minutes is:

$$P(X=x|t=2)=\frac{5^x e^{-5}}{x!}\ ,\quad x=0,1,2,\ldots\rightarrow$$

$$P(X=2|t=2)=\frac{5^2 e^{-5}}{2!}$$

$$=\frac{25}{2}e^{-5}$$

$$=0.08422$$

3. The probability that at least two will arrive within 3 minutes.

Solution:
The probability that at least two will arrive within 3 minutes is:

$$P(X=x|t=3)=\frac{7.5^x e^{-7.5}}{x!}\ ,\quad x=0,1,2,\ldots\rightarrow$$

$$P(X\geq 2|t=3)=1-P(X<2)$$

$$=1-[P(X=0)+P(X=1)]$$

$$=1-\left[\frac{7.5^0 e^{-7.5}}{0!}+\frac{7.5^1 e^{-7.5}}{1!}\right]$$

$$=1-e^{-7.5}[1+7.5]$$

$$=1-(8.5)(0.00055)$$

$$=0.9953$$

4. The standard deviation of the number of customers who will arrive to the counter within one and a half hours.

Solution:
The standard deviation of the number of customers who will arrive to the counter within one and a half hours is given by:

$$\sigma = \sqrt{\lambda t}$$
$$= \sqrt{(2.5)(90)}$$
$$= \sqrt{225} = 15 \ customers$$

Recap

In this lecture, you have learned that:

- If the random variable X counts the number (x) of successes from (K) successes in the population in a sample of size (n) drawn from a population of size (N), then X is said to have the Hypergeometric Probability Distribution

- There are four characteristics of the Hypergeometric Distribution
  - The n trials in a sample are taken from a finite population of size N
  - The sample is taken without replacement
  - The trials are dependent
  - The probability of success is NOT fixed