# Module 4. Special distribution

| Discrete distributions | Continuous distributions |
|---|---|
| 1. Binomial distribution | 1.Uniform distribution (or) Rectangular distribution |
| 2. Poisson distribution | 2.Exponential distribution |
| 3. Geometric distribution | 3.Gamma distribution |
| 4.Negative binomial distribution | 4.Weibull distribution |
| | 5.Normal distribution |

## Properties:

1. There must be a fixed number of trials.
2. All trials must have identical probabilities of success (P).
3. The trials must be independent of each other.

## Notations:

n = number of trials

p = probability of success

q = probability of failure

X = A random variable which represents the number of successes.

## Applications:

**Sampling process (or) Quality control measures in industries to classify the items as defective or non-defective.**

## Binomial distribution:

### Probability mass function of Binomial distribution:

**A discrete random variable X is said to follow binomial distribution if its probability mass function is**

$$P(x) = n_{c_x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots\ldots n$$

### Moment generating function:

**The moment generating function of a binomial variate is**

$$M_x(t) = E[e^{tx}]$$

$$= \sum_{x=0}^{n} e^{tx} p(x)$$

$$= \sum_{x=0}^{n} e^{tx} n_{c_x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} n_{c_x} (pe^t)^x q^{n-x}$$

$$[(x+a)^n = n_{c_0} x^n + n_{c_1} x^{n-1} a + n_{c_2} x^{n-2} a^2 + \ldots\ldots\ldots\ldots + n_{c_n} a^n]$$

$$= (Pe^t + q)^n$$

**Mean and Variance of Binomial distribution:**

$$\mu_1^1 = [\frac{d}{dt}\{M_x(t)\}]_{t=0}$$

$$= [\frac{d}{dt}\{(pe^t + q)^n\}]_{t=0}$$

$$= [n(pe^t + q)^{n-1} pe^t]_{t=0}$$

$$= n(p+q)^{n-1} p \qquad [\because p+q=1]$$

$$= np$$

**Mean of the Binomial distribution is = np**

$$\mu_2^1 = [\frac{d}{dt^2}^2 \{M_x(t)\}]_{t=0}$$

$$= [\frac{d^2}{dt^2}\{(pe^t + q)^n\}]_{t=0}$$

$$= [\frac{d}{dt}\{n(pe^t + q)^{n-1} pe^t\}]_{t=0}$$

$$= np[e^t(n-1)(pe^t + q)^{n-2} pe^t + (pe^t + q)^{n-1} e^t]_{t=0}$$

$$= np[(n-1)(p+q)^{n-2} p + (p+q)^{n-1}] \qquad [\because p+q=1]$$

$$= np[(n-1)p+1]$$

$$\mu_2^1 = n(n-1)p^2 + np$$

**Variance of Binomial distribution:**

$$\mu_2 = \mu_2^1 - (\mu_1^1)^2$$
$$= n(n-1)p^2 + np - n^2p^2$$
$$= n^2p^2 - np^2 + np - n^2p^2$$
$$= np(1-p)$$
$$= npq$$

**Variance of Binomial distribution is npq**

**Problems:**

1. **The mean of a binomial distribution is 5 and standard deviation is 2. Determine the distribution.**

**Given mean = np = 5**

**S.D** $= \sqrt{npq} = 2$

**npq = 4**

**5q = 4**

**q = 4/5**

**p = 1-q=1/5**

**We have np = 5**

**n(1/5) = 5**

**n = 25**

**Hence the Binomial distribution is**

$$p(X=x) = n_{C_x} p^x q^{n-x} = 25_{C_x} (\frac{1}{5})^x (\frac{4}{5})^{25-x}, \quad x = 0,1,2\ldots\ldots 25.$$

2. **The mean and variance of a binomial distribution are 4 and 4/3. Find** $P(X \geq 1)$

**Given mean np = 4**

**Variance npq = 4/3**

$$\frac{npq}{np} = \frac{4/3}{4} = \frac{1}{3}$$

**q=1/3 ; p=2/3**

**np = 4**

**n(2/3) = 4**

**n = 6**

$$p(X \geq 1) = 1 - P(X < 1)$$

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

$$= 1 - 6_{C_0} p^0 q^{6-0}$$

$$= 1 - 6_{C_0} (2/3)^0 (1/3)^{6-0}$$

$$= 1 - (1/3)^6 = 0.998$$

**3. An irregular 6 faced die is such that the probability that it gives 3 even numbers in 5 throws is twice the probability that it gives 2 even numbers in 5 throws. How many sets of exactly 5 trials can be expected to give no even number out of 2500 sets?**

Let X denote the number of even numbers obtained in 5 trias.

Given P(x=3)=2 *P(x=2)

$$5_{C_3}\, p^3 q^2 = 2 * 5_{C_2}\, p^2 q^3$$

$$p = 2q = 2(1 - p) = 2 - 2p$$

$$3p = 2$$

i.e., $p = 2/3$

$$q = 1/3$$

Now, P(getting no even number)=P(x=0)

$$5_{C_0}\, p^0 q^5 = (1/3)^5 = \frac{1}{243}$$

Number of sets having no success (no even number) out of N sets=N *P(x=0)

Required number of sets=2500*1/243=10, nearly.

## Poisson distribution

A discrete random variable X is said to follow Poisson distribution with parameter $\lambda$ if its probability mass function is $P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0,1,2.........\infty$

Poisson distribution as limiting form of Binomial distribution under the following conditions

i) n, the number of trials is indefinitely large, i.e $n \to \infty$.

ii) p, the constant probability of success in each trial is very small. i.e., $p \to 0$.

iii) np ($=\lambda$) is finite (or) $p=\lambda/n$ and $q=1-\lambda/n$, where $\lambda$ is a positive real number.

The following are some of the examples, which may be analysed using Poisson distribution.

1. The number of alpha particles emitted by a radioactive source in a given time interval.
2. The number of telephone calls received at a telephone exchange in a given time interval.

3. The number of defective articles in a packet of 100.
4. The number of printing errors at each page of a book.
5. The number of road accidents reported in a city per day.

**Moment generating function:**

The MGF of the Poisson variate X with parameter $\lambda$ is given as

$$M_x(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^t} \qquad \because \sum_{x=0}^{\infty} \frac{a^x}{x!} = e^a$$

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

## Mean and variance of Poisson distribution

$$\mu_1^1 = \left[\frac{d}{dt} M_x(t)\right]_{t=0}$$

$$= \left[\frac{d}{dt} e^{-\lambda} e^{\lambda e^t}\right]_{t=0}$$

$$= \left[e^{-\lambda} e^{\lambda e^t} \lambda e^t\right]_{t=0}$$

$$= e^{-\lambda} e^{\lambda} \lambda$$

$$= \lambda$$

Mean of Poisson distribution = $\lambda$

$$\mu_2^1 = \left[\frac{d^2}{dt^2} M_x(t)\right]_{t=0}$$

$$= \left[\frac{d}{dt}\left(\frac{d}{dt} e^{-\lambda} e^{\lambda e^t}\right)\right]_{t=0}$$

$$= \left[\frac{d}{dt}\left(e^{-\lambda} e^{\lambda e^t} \lambda e^t\right)\right]_{t=0}$$

$$= \lambda e^{-\lambda} [e^{\lambda e^t} . e^t + e^t e^{\lambda e^t} \lambda e^t]_{t=0}$$

$$= \lambda e^{-\lambda} [e^{\lambda} + e^{\lambda} \lambda]$$

$$= \lambda e^{-\lambda} e^{\lambda} + \lambda^2 e^{-\lambda} e^{\lambda}$$

$$\mu_2^1 = \lambda^2 + \lambda$$

## Variance of Poisson distribution

$$\mu_2 = \mu_2^1 - (\mu_1^1)^2$$
$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

In Poisson distribution Mean = Variance = $\lambda$

## Problems:

1. The number of monthly breakdowns of a computer is a random variable having a Poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month.

   a) Without a breakdown
   b) With only one breakdown and
   c) With atleast one breakdown

Let X denotes the number of breakdowns of the computer in a month.

X follows a Poisson distribution with mean $\lambda = 1.8$

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$
$$= \frac{e^{-1.8} (1.8)^x}{x!}$$

$a)\ p(x=0)=e^{-1.8}=0.1653$

$b)\ p(x=1)=e^{-1.8}(1.8)=0.2975$

$c)\ p(x\geq 1)=1-p(x=0)=0.8347$

**2.** If the probability of a defective fuse from a manufacturing unit is 2%, in a box of 200 fuses, find the probability that a) exactly 4 fuses are defective b) more than 3 fuses are defective.

P=0.02, n=200

Mean = $\lambda$ =np=200*0.02=4

$$a)\ P(X=4)=\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$=\frac{e^{-4}(4)^{4}}{4!}=0.1952$$

$$b)\ P(x>3)=1-P(x\leq 3)$$

$$=1-[p(x=3)+p(x=2)+p(x=1)+p(x=0)]$$

$$=1-e^{-4}[\frac{4^3}{3!}+\frac{4^2}{2!}+\frac{4^1}{1!}+\frac{4^0}{0!}]=0.5669$$

## Fitting of Binomial Distribution

Fitting a binomial distribution means assuming that the given distribution is approximately binomial and hence finding the probability mass function and then finding the theoretical frequencies.

**Problem:**

**Fit a binomial distribution for the following data:**

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|---|
| f | 5 | 18 | 28 | 12 | 7 | 6 | 4 | 80 |

| x | f | fx | Theoretical frequency |
|---|---|---|---|
| 0 | 5 | 0 | 4 |
| 1 | 18 | 18 | 15 |
| 2 | 28 | 56 | 25 |
| 3 | 12 | 36 | 22 |
| 4 | 7 | 28 | 11 |
| 5 | 6 | 30 | 3 |
| 6 | 4 | 24 | 0 |
| | $\sum f = N = 80$ | $\sum fx = 192$ | |

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{192}{80} = 2.4$$

$$np = 2.4 (or) 6p = 2.4$$

$$p = 0.4, \quad q = 0.6$$

**Theoretical distribution is given by**

$$= N \times \quad P(X = x)$$

$$= 80 \quad \times \sum_{x=0}^{6} 6_{c_x} p^x q^{n-x}$$

$$= 80 \times q^6 + 6_{c_1} q^{6-1} p + 6_{c_2} q^{6-2} p^2 + \ldots\ldots$$

$$= 80 \times [(0.6)^6 + 6_{c_1} (0.6)^{6-1} (0.4) + 6_{C_2} (0.6)^{6-2} (0.4)^2 + 6_{c_3} (0.6)^{6-3} (0.4)^3$$
$$+ 6_{c_4} (0.6)^{6-4} (0.4)^4 + 6_{c_5} (0.6)^{6-5} (0.4)^5 + 6_{c_6} (0.4)^6]$$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$: | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| $f$: | 3.73 | 14.93 | 24.88 | 22.12 | 11.06 | 2.95 | 0.33 |

## Fitting of Poisson Distribution

1. Fit a Poisson distribution to the following data and calculate the theoretical frequencies.

| **Deaths** | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| **Frequency** | 122 | 60 | 15 | 2 | 1 |

| **x** | **f** | **fx** | **Theoretical frequency** |
|---|---|---|---|
| **0** | **122** | **0** | **121** |
| **1** | **60** | **60** | **61** |
| **2** | **15** | **30** | **15** |
| **3** | **2** | **6** | **3** |
| **4** | **1** | **4** | **0** |
| | $N = 200$ | $\sum fx = 100$ | |

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{100}{200} = 0.5$$

**Theoretical distribution is given by**

$$= N \times \quad P(X = x)$$

$$= 200 \quad \times \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= 200 \times \frac{e^{-0.5} (0.5)^x}{x!} \quad \textit{putting} \quad x = 0, \ 1, \ 2, \ 3, \ 4$$

| | | | | | |
|---|---|---|---|---|---|
| $x$: | 0 | 1 | 2 | 3 | 4 |
| $f$: | 121 | 61 | 15 | 3 | 0 |

---

## Exponential Distribution:

**Definition:** A continuous random variable $X$ is said to follow an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$$f(x|\lambda) = \lambda e^{-\lambda x}, \ x > 0, \lambda > 0.$$

It is also known as negative exponential distribution.

- Mean of the exponential distribution is $1/_{\lambda}$.
- Variance of the exponential distribution is $1/_{\lambda^2}$ .

- The moment generating function of the exponential distribution is $\lambda/_{\lambda - t}$ , $\lambda > t$.
- The $r^{\text{th}}$ moment about the origin is $r!/_{\lambda^r}$ , $r = 1, 2, 3, \ . \ . \ ..$

**Examples related to exponential distribution:**

## Example:

The mileage which car owners get with a certain kind of radial tire is a RV having an exponential distribution with mean 40,000 km. Find the probabilities that one of these tires will last (i) at least 20,000 km and (ii) at most 30,000 km.

**Solution:**

Let $X$ denote the mileage obtained with the tire

$$f(x) = \frac{1}{40,000} e^{-x/40,000} \quad x > 0$$

$$\text{(i) } P(X \geq 20,000) = \int_{20,000}^{\infty} \frac{1}{40,000} e^{-x/40,000} \, dx$$

$$= \left[-e^{-x/40,000}\right]_{20,000}^{\infty}$$

$$= e^{-0.5} = 0.6065$$

$$\text{(ii) } P(X \leq 30,000) = \int_{0}^{30,000} \frac{1}{40,000} e^{-x/40,000} \, dx$$

$$= \left[-e^{-x/40,000}\right]_{0}^{30,000}$$

$$= 1 - e^{-0.75} = 0.5270$$

**Exercise:**

The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = 1/2$.

(a) What is the probability that the repair time exceeds 2 h?

(b) What is the conditional probability that a repair takes at least 10 h given that its duration exceeds 9 h?

HINT:

$X \sim ED(\lambda)$, with $\lambda = 1/2$

(a) $P(X > 2)$

(b) $P(X \geq 10 / X > 9) = \dfrac{P(X \geq 10, X > 9)}{P(X > 9)}$

## Normal Distribution:

**Definition:** A continuous random variable X is said to have a normal distribution if its probability function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \ \sigma > 0,$$

where $\mu$ and $\sigma$ are the mean and standard deviation of the normal random variable $X$, respectively.

If $X$ follows normal distribution, then it is denoted by $X \sim N(\mu, \sigma)$. It is also called as the Gaussian distribution.

# Standard Normal Distribution:

**Definition:** The standard normal random variable $Z = \frac{X-\mu}{\sigma}$ is said to have a standard normal distribution if its probability function is defined by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty.$$

A standard normal variate is denoted by $N(0,1)$. The standard normal distribution is also known as a Z-distribution or a unit normal distribution. Standardization of a normal distribution helps us to make use of the tables of the area of the standard curve.

# Distribution function of a standard normal variable:

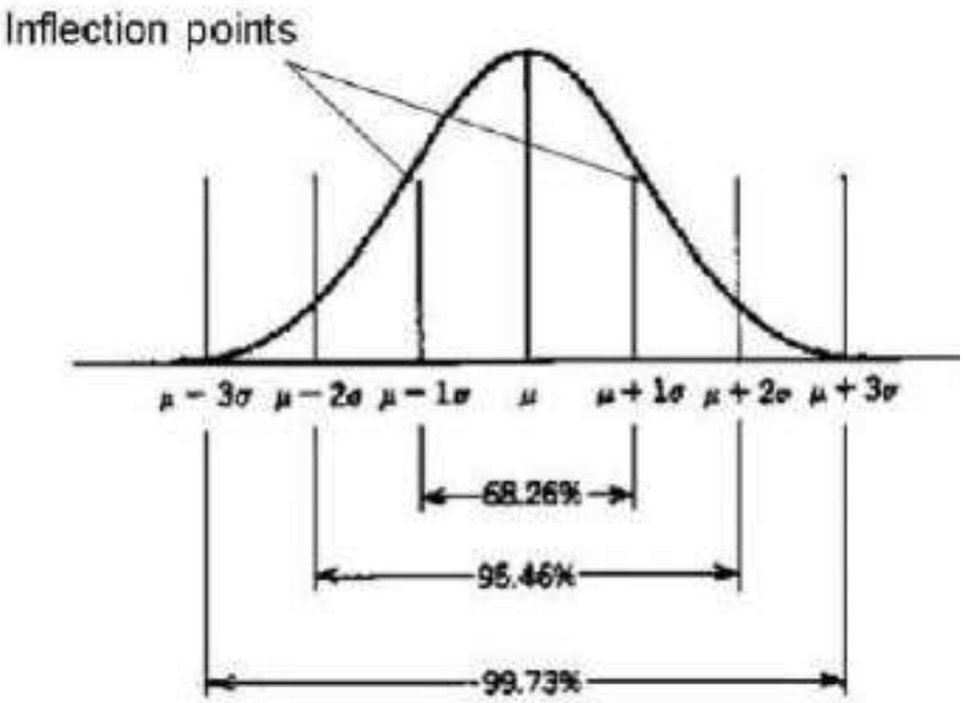
The distribution function $\phi(z)$ of a standard normal variate, $Z$ is defined by

$$\begin{aligned} \phi(z) = P(Z \le z) &= \int_{-\infty}^{z} f(t)dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} dt. \end{aligned}$$

**Properties of $\phi(z)$:**

1. $\phi(-z) = 1 - \phi(z)$
2. $P(a \le X \le b) = \phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma})$
3. $P(Z \le a) = P(Z \ge -a)$

## Graph of normal distribution:




## Area under normal curve:

The curve of a normal distribution is bell shaped, with the highest point over the mean $\mu$. It is symmetrical about a vertical line through $\mu$. The normal curve has the following features.

1. The curve is symmetrical about the coordinate at its mean which locates the peak of the bell.
2. The values of mean, median and mode are equal.
3. The curve extends from $-\infty$ to $\infty$.
4. The area covered between $\mu - \sigma$ and $\mu + \sigma$ is 0.6826 (i.e. 68.26 % of area).
5. The area covered between $\mu - 2\sigma$ and $\mu + 2\sigma$ is 0.9544 (i.e. 95.44 % of area).
6. The area covered between $\mu - 3\sigma$ and $\mu + 3\sigma$ is 0.9974 (i.e. 99.74 % of area).

## AREAS UNDER NORMAL CURVE

Normal probability curve is given by

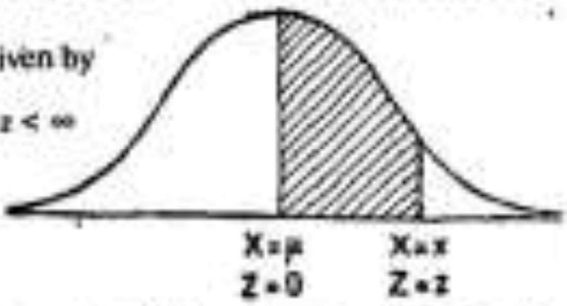
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad -\infty < x < \infty$$

and standard normal probability curve is given by

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right), \; -\infty < z < \infty$$

where $Z = \dfrac{X - E(X)}{\sigma_X} \sim N(0, 1)$

Areas under Normal Curve




The following table gives the shaded area in the diagram viz., $P(0 < Z < z)$ for different values of $z$.

### TABLE OF AREAS

| $z \rightarrow$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| ·0 | ·0000 | ·0040 | ·0080 | ·0120 | ·0160 | ·0199 | ·0239 | ·0279 | ·0319 | ·0359 |
| ·1 | ·0398 | ·0438 | ·0478 | ·0517 | ·0557 | ·0596 | ·0636 | ·0675 | ·0714 | ·0753 |
| ·2 | ·0793 | ·0832 | ·0871 | ·0910 | ·0948 | ·0987 | ·1026 | ·1064 | ·1103 | ·1141 |
| ·3 | ·1179 | ·1217 | ·1255 | ·1293 | ·1331 | ·1368 | ·1406 | ·1443 | ·1480 | ·1517 |
| ·4 | ·1554 | ·1591 | ·1628 | ·1664 | ·1700 | ·1736 | ·1772 | ·1808 | ·1844 | ·1879 |
| ·5 | ·1915 | ·1950 | ·1985 | ·2019 | ·2054 | ·2088 | ·2123 | ·2157 | ·2190 | ·2224 |
| ·6 | ·2257 | ·2291 | ·2324 | ·2357 | ·2389 | ·2422 | ·2454 | ·2486 | ·2517 | ·2549 |
| ·7 | ·2580 | ·2611 | ·2642 | ·2673 | ·2703 | ·2734 | ·2764 | ·2794 | ·2823 | ·2852 |
| ·8 | ·2881 | ·2910 | ·2939 | ·2967 | ·2995 | ·3023 | ·3051 | ·3078 | ·3106 | ·3133 |
| ·9 | ·3159 | ·3186 | ·3212 | ·3238 | ·3264 | ·3289 | ·3315 | ·3340 | ·3365 | ·3389 |
| 1·0 | ·3413 | ·3438 | ·3461 | ·3485 | ·3508 | ·3531 | ·3554 | ·3577 | ·3599 | ·3621 |
| 1·1 | ·3643 | ·3665 | ·3686 | ·3708 | ·3729 | ·3749 | ·3770 | ·3790 | ·3810 | ·3830 |
| 1·2 | ·3849 | ·3869 | ·3888 | ·3907 | ·3925 | ·3944 | ·3962 | ·3980 | ·3997 | ·4015 |
| 1·3 | ·4032 | ·4049 | ·4066 | ·4082 | ·4099 | ·4115 | ·4131 | ·4147 | ·4162 | ·4177 |
| 1·4 | ·4192 | ·4207 | ·4222 | ·4236 | ·4251 | ·4265 | ·4279 | ·4292 | ·4306 | ·4319 |
| 1·5 | ·4332 | ·4345 | ·4357 | ·4370 | ·4382 | ·4394 | ·4406 | ·4418 | ·4429 | ·4441 |
| 1·6 | ·4452 | ·4463 | ·4474 | ·4484 | ·4495 | ·4505 | ·4515 | ·4525 | ·4535 | ·4545 |
| 1·7 | ·4554 | ·4564 | ·4573 | ·4582 | ·4591 | ·4599 | ·4608 | ·4616 | ·4625 | ·4633 |
| 1·8 | ·4641 | ·4649 | ·4656 | ·4664 | ·4671 | ·4678 | ·4686 | ·4693 | ·4699 | ·4706 |
| 1·9 | ·4713 | ·4719 | ·4726 | ·4732 | ·4738 | ·4744 | ·4750 | ·4756 | ·4761 | ·4767 |
| 2·0 | ·4772 | ·4778 | ·4783 | ·4788 | ·4793 | ·4798 | ·4803 | ·4808 | ·4812 | ·4817 |
| 2·1 | ·4821 | ·4826 | ·4830 | ·4834 | ·4838 | ·4842 | ·4846 | ·4850 | ·4854 | ·4857 |
| 2·2 | ·4861 | ·4864 | ·4868 | ·4871 | ·4875 | ·4878 | ·4881 | ·4884 | ·4887 | ·4890 |
| 2·3 | ·4893 | ·4896 | ·4898 | ·4901 | ·4904 | ·4906 | ·4909 | ·4911 | ·4913 | ·4916 |
| 2·4 | ·4918 | ·4920 | ·4922 | ·4925 | ·4927 | ·4929 | ·4931 | ·4932 | ·4934 | ·4936 |
| 2·5 | ·4938 | ·4940 | ·4941 | ·4943 | ·4945 | ·4946 | ·4948 | ·4949 | ·4951 | ·4952 |
| 2·6 | ·4953 | ·4955 | ·4956 | ·4957 | ·4959 | ·1960 | ·4961 | ·4962 | ·4963 | ·4964 |
| 2·7 | ·4965 | ·4966 | ·4967 | ·4968 | ·4969 | ·4970 | ·4971 | ·4972 | ·4973 | ·4974 |
| 2·8 | ·4974 | ·4975 | ·4976 | ·4977 | ·4977 | ·4978 | ·4979 | ·4979 | ·4980 | ·4981 |
| 2·9 | ·4981 | ·4982 | ·4982 | ·4983 | ·4984 | ·4984 | ·4985 | ·4985 | ·4986 | ·4986 |
| 3·0 | ·4987 | ·4987 | ·4987 | ·4988 | ·4988 | ·4989 | ·4989 | ·4989 | ·4990 | ·4990 |
| 3·1 | ·4990 | ·4991 | ·4991 | ·4991 | ·4992 | ·4992 | ·4992 | ·4992 | ·4993 | ·4993 |
| 3·2 | ·4993 | ·4993 | ·4994 | ·4994 | ·4994 | ·4994 | ·4994 | ·4995 | ·4995 | ·4995 |
| 3·3 | ·4995 | ·4995 | ·4995 | ·4996 | ·4996 | ·4996 | ·4996 | ·4996 | ·4996 | ·4997 |
| 3·4 | ·4997 | ·4997 | ·4997 | ·4997 | ·4997 | ·4997 | ·4997 | ·4997 | ·4997 | ·4998 |
| 3·5 | ·4998 | ·4998 | ·4998 | ·4998 | ·4998 | ·4998 | ·4998 | ·4998 | ·4998 | ·4998 |
| 3·6 | ·4998 | ·4998 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 |
| 3·7 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 | ·4999 |
| 3·9 | ·5000 | ·5000 | ·5000 | ·5000 | ·5000 | ·5000 | ·5000 | ·5000 | ·5000 | ·5000 |

## Area under standard normal curve:

The standard normal random variable $Z$ and the normal random variable $X$ have identical curves.

1. The standard normal curve covers 68.26 % area between $z = -1$ and $z = 1$. Therefore, $P(-1 \leq Z \leq 1) = 0.6826$.
2. The standard normal curve covers 95.44 % area between $z = -2$ and $z = 2$. Therefore, $P(-2 \leq Z \leq 2) = 0.9544$.
3. The standard normal curve covers 99.74 % area between $z = -3$ and $z = 3$. Therefore, $P(-3 \leq Z \leq 3) = 0.9974$.

## Properties of a normal curve:

- In a normal distribution, mean deviation about mean is approximately equal to $\frac{4}{5}$ times its standard deviation.
- In a normal distribution, the quartiles $Q_1$ and $Q_3$ are equidistant from the median.
- The normal curve is bell shaped and is symmetrical about $X = \mu$.
- The area bound by the normal curve and the $x$-axis equal to 1.
- The tails of the curve of a normal distribution extend identically in both sides of $x = \mu$ and never touch the $x$-axis.
- The moment generating function of a normal distribution with respect to origin is $e^{\mu t + \frac{1}{2}t^2\sigma^2}$.

# Examples related to normal distribution:

## Example:

The marks obtained by a number of students in a certain subject are approximately normally distributed with mean 65 and standard deviation 5. If 3 students are selected at random from this group, what is the probability that at least 1 of them would have scored above 75?

## Solution:

If $X$ represents the marks obtained by the students, $X$ follows the distribution $N(65, 5)$.

$P$(a student scores above 75)

$$= P(X > 75) = P\left(\frac{75-65}{5} < \frac{X-65}{5} < \infty\right)$$
$$= P(2 < Z < \infty), \text{ (where } Z \text{ is the standard normal variate)}$$
$$= 0.5 - P(0 < Z < 2)$$
$$= 0.5 - 0.4772, \text{ (from the table of areas)}$$
$$= 0.0228$$

Let $p = P$(a student scores above 75) $= 0.0228$ then $q = 0.9772$ and $n = 3$.

Since $p$ is the same for all the students, the number $Y$, of (successes) students scoring above 75, follows a binomial distribution.

$P$(at least 1 student scores above 75)

$$= P(\text{at least 1 success})$$
$$= P(Y \geq 1) = 1 - P(Y = 0)$$
$$= 1 - nC_0 \times P^0 q^n$$
$$= 1 - 3C_0\,(0.9772)^3$$
$$= 1 - 0.9333$$
$$= 0.0667$$

## Example:

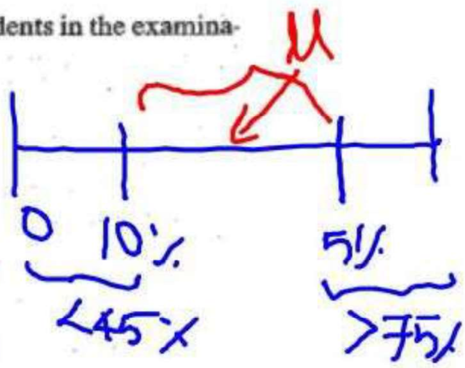
In an engineering examination, a student is considered to have failed, secured second class, first class and distinction, according as he scores less than 45%, between 45% and 60%, between 60% and 75% and above 75% respectively. In a particular year 10% of the students failed in the examination and 5% of the students got distinction. Find the percentages of students who have got first class and second class. (Assume normal distribution of marks).

## Solution:

Let $X$ represent the percentage of marks scored by the students in the examination.

Let $X$ follow the distribution $N(\mu, \sigma)$.

**Given:** $P(X < 45) = 0.10$ and $P(X > 75) = 0.05$




i.e., $$P\left(-\infty < \frac{X-u}{\sigma} < \frac{45-\mu}{\sigma}\right) = 0.10 \text{ and}$$

$$P\left(\frac{75-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \infty\right) = 0.05$$

-ve

i.e., $$P\left(-\infty < Z < \frac{45-\mu}{\sigma}\right) = 0.10 \text{ and}$$

+ve

$$P\left(\frac{75-\mu}{\sigma} < Z < \infty\right) = 0.05$$

$45 < \mu < 75$

$\therefore$ $$P\left(0 < Z < \frac{\mu-45}{\sigma}\right) = 0.40 \text{ and}$$

$$P\left(0 < Z < \frac{75-\mu}{\sigma}\right) = 0.45$$

From the table of areas, we get

$$\frac{\mu-45}{\sigma} = 1.28 \text{ and } \frac{75-\mu}{\sigma} = 1.64$$

i.e., $\mu - 1.28\,\sigma = 45$ (1)

and $\mu + 1.64\,\sigma = 75$ (2)

Solving equations (1) and (2), we get

$\mu = 58.15$ and $\sigma = 10.28$

Now $P$ (a student gets first class)

$$= P(60 < X < 75)$$

$$= P\left\{\frac{60-58.15}{10.28} < Z < \frac{75-58.15}{10.28}\right\}$$

$$= P\{0.18 < Z < 1.64\}$$

$$= P\{0 < Z < 1.64\} - P\{0 < Z < 0.18\}$$

$$= 0.4495 - 0.0714 = 0.3781$$

$\therefore$ Percentage of students getting first class = 38 (approximately)

Now percentage of students getting second class

= 100 – (sum of the percentages of students who have failed, got first class and got distinction)

= 100 – (10 + 38 + 5), approximately.

= 47 (approximately)

## WEIBULL DISTRIBUTION

**Definition:** A continuous RV $X$ is said to follow a *Weibull distribution* with parameters $\alpha, \beta > 0$, if the RV $Y = \alpha X^{\beta}$ follows the exponential distribution with density function $f_Y(y) = e^{-y}, y > 0$.

### Density Function of the Weibull Distribution

Since $Y = \alpha \cdot X^{\beta}$, we have $y = \alpha \cdot x^{\beta}$.

By the transformation rule, derived in chapter 3, we have $f_X(x) = f_Y(y)\left|\dfrac{dy}{dx}\right|$,

where $f_X(x)$ and $f_Y(y)$ are the density functions of $X$ and $Y$ respectively.

$$\therefore \quad f_X(x) = e^{-y}\, \alpha\beta x^{\beta-1}$$
$$= \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}};\ x > 0 \quad [\because y > 0]$$

### Mean and Variance of Weibull Distribution

The raw moments $\mu_r'$ about the origin of the Weibull distribution are given by

$$\mu_r' = E(X^r)$$
$$= \alpha\beta \int_0^{\infty} x^{r+\beta-1} e^{-\alpha x^{\beta}}\, dx$$
$$= \int_0^{\infty} \left(\frac{y}{\alpha}\right)^{\frac{r}{\beta}+1-\frac{1}{\beta}} e^{-y} \left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}-1} dy,$$

on putting $\quad y = \alpha x^{\beta}$ or $x = \left(\dfrac{y}{\alpha}\right)^{\frac{1}{\beta}}$

$$= \alpha^{-r/\beta} \int_0^{\infty} y^{r/\beta} e^{-y}\, dy$$
$$= \alpha^{-r/\beta}\, \Gamma\left(\frac{r}{\beta}+1\right)$$

$$\therefore \quad \text{Mean} = E(X) = \mu_1' = \alpha^{-\frac{1}{\beta}}\, \Gamma\left(\frac{1}{\beta}+1\right)$$
$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$
$$= \alpha^{-2/\beta}\left[\Gamma\left(\frac{2}{\beta}+1\right) - \left\{\Gamma\left(\frac{1}{\beta}+1\right)\right\}^2\right]$$

## PROBLEM 1

In a certain city, the daily consumption of electric power in millions of kilowatt-hours can be treated as a RV having an Erlang distribution with parameters $\lambda = \frac{1}{2}$ and $k = 3$. If the power plant of this city has a daily capacity of 12 millions kilowatt-hours, what is the probability that this power supply will be inadequate on any given day.

Let $X$ represent the daily consumption of electric power (in millions of kilo-watt-hours). Then the density function of $X$ is given as

$$f(x) = \frac{\left(\frac{1}{2}\right)^3}{\Gamma(3)} x^2 e^{-x/2}, x > 0$$

$P$(the power supply is inadequate)

$$= P(X > 12) = \int_{12}^{\infty} f(x)\, dx \quad [\because \text{The daily capacity is only } 12]$$

$$= \int_{12}^{\infty} \frac{1}{\Gamma(3)} \cdot \frac{1}{2^3} x^2 e^{-x/2}\, dx$$

$$= \frac{1}{16}\left[x^2\left(\frac{e^{-x/2}}{-\frac{1}{2}}\right) - 2x\left(\frac{e^{-x/2}}{\frac{1}{4}}\right) + 2\left(\frac{e^{-x/2}}{-\frac{1}{8}}\right)\right]_{12}^{\infty}$$

$$= \frac{1}{16} e^{-6} (288 + 96 + 16)$$

$$= 25\, e^{-6} = 0.0625$$

If a company employes $n$ sales persons, its gross sales in thousands of rupees may be regarded as a RV having an Erlang distribution with $\lambda = \frac{1}{2}$ and $k = 80\sqrt{n}$. If the sales cost is Rs. 8000 per salesperson, how many salespersons should the company employ to maximise the expected profit?

Let $X$ represent the gross sales (in Rupees) by $n$ salespersons.

$X$ follows the Erlang distribution with parameters $\lambda = \frac{1}{2}$ and $k = 80{,}000\sqrt{n}$.

$$\therefore\ E(X) = \frac{k}{\lambda} = 1{,}60{,}000\sqrt{n}$$

If y denotes the total expected profit of the company, then

$$y = \text{total expected sales} - \text{total sales cost}$$

$$= 1{,}60{,}000\sqrt{n} - 8000\,n$$

$$\frac{dy}{dn} = \frac{80{,}000}{\sqrt{n}} - 8000$$

$$= 0, \text{ when } \sqrt{n} = 10 \text{ or } n = 100$$

$$\frac{d^2y}{dn^2} = -\frac{40{,}000}{n^{3/2}} < 0, \text{ when } n = 100.$$

Therefore, $y$ is maximum, when $n = 100$.

That is the company should employ 100 salespersons in order to maximise the total expected profit.

Consumer demand for milk in a certain locality, per month, is known to be a general Gamma (Erlang) RV. If the average demand is $a$ litres and the most likely demand is $b$ litres ($b < a$), what is the variance of the demand?

Let $X$ represent the monthly consumer demand of milk.

Average demand is the value of $E(X)$.

Most likely demand is the value of the mode of $X$ or the value of $X$ for which its density function is maximum.

If $f(x)$ is the density function of $X$, then

$$f(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} \quad x > 0$$

$$f'(x) = \frac{\lambda^k}{\Gamma(k)} [(k-1) x^{k-2} e^{-\lambda x} - \lambda x^{k-1} e^{-\lambda x}]$$

$$= \frac{\lambda^k}{\Gamma(k)} x^{k-2} e^{-\lambda x} \{(k-1) - \lambda x\}$$

$$= 0, \text{ when } x = 0, x = \frac{k-1}{\lambda}$$

$$f''(x) = \frac{\lambda^k}{\Gamma(k)} [-\lambda x^{k-2} e^{-\lambda x} + \{(k-1) - \lambda x\} \frac{d}{dx} \{x^{k-2} e^{-\lambda x}\}]$$

$$< 0, \text{ when } x = \frac{k-1}{\lambda}$$

Therefore $f(x)$ is maximum, when $x = \dfrac{k-1}{\lambda}$.

i.e., Most likely demand $= \dfrac{k-1}{\lambda} = b$ (1)

and $E(X) = \dfrac{k}{\lambda} = a$ (2)

Now $\text{Var}(X) = \dfrac{k}{\lambda^2} = \dfrac{k}{\lambda} \cdot \dfrac{1}{\lambda}$

$= a(a-b)$, [from (1) and (2)]

Each of the 6 tubes of a radio set has a life length (in years) which may be considered as a RV that follows a Weibull distribution with parameters $\alpha = 25$ and $\beta = 2$. If these tubes function independently of one another, what is the probability that no tube will have to be replaced during the first 2 months of service?

If $X$ represents the life length of each tube, then its density function $f(x)$ is given by

$$f(x) = \alpha\beta\, x^{\beta-1}\, e^{-\alpha x^{\beta}} \quad x > 0$$

i.e., $f(x) = 50x\, e^{-25x} \quad x > 0$

Now $P$(a tube is not to be replaced during the first 2 months)

$$= P\left(X > \frac{1}{6}\right)$$

$$= \int_{\frac{1}{6}}^{\infty} 50x\, e^{-25x^2}\, dx$$

$$= \left(-e^{-25x^2}\right)_{\frac{1}{6}}^{\infty} = e^{-25/36}$$

. $P$(all the 6 tubes are not to be replaced during the first 2 months)

$$= (e^{-25/36})^6 \quad \text{(by independence)}$$

$$= e^{-25/6}$$

$$= 0.0155$$

If the time $T$ to failure of a component follows a Weibull distribution with parameters $\alpha$ and $\beta$, find the hazard rate or conditional failure rate at time $t$ of the component.

Refer to Example (19) in Worked Example 2(A).

If $f(t)$ is the density function of $T$ and $h(t)$ is the hazard rate at time $t$, then

$$h(t) = \frac{f(t)}{1 - F(t)}$$

where $F(t)$ is the distribution function of $T$.

Now $\quad f(t) = \alpha\beta \times t^{\beta - 1} e^{-\alpha t^\beta} \quad t > 0$

$\therefore \quad F(t) = P(T \leq t)$

$$= \int_0^t \alpha\beta t^{\beta - 1} \cdot e^{-\alpha t^\beta} \, dt$$

$$= \left[ -e^{-\alpha t^\beta} \right]_0^t$$

$$= 1 - e^{-\alpha t^\beta}$$

$$\therefore \quad h(t) = \frac{\alpha\beta t^{\beta-1} e^{-\alpha t^\beta}}{e^{-\alpha t^\beta}}$$

$$= \alpha\beta t^{\beta - 1}$$

If the life $X$ (in years) of a certain type of car has a Weibull distribution with the parameter $\beta = 2$, find the value of the parameter $\alpha$, given that probability that the life of the car exceeds 5 years is $e^{-0.25}$. For these values of $\alpha$ and $\beta$, find the mean and variance of $X$.

The density function of $X$ is given by

$$f(x) = 2\alpha x\, e^{-\alpha x^2},\ x > 0 \quad [\because \beta = 2]$$

Now $P(X > 5) = \int_5^\infty 2\alpha x\, e^{-\alpha x^2}\, dx$

$$= \left(-e^{-\alpha x^2}\right)_5^\infty$$

$$= e^{-25\alpha}$$

Given that $P(X > 5) = e^{-0.25}$

$\therefore \quad e^{-25\alpha} = e^{-0.25}$

$\therefore \quad \alpha = \frac{1}{100}$

For the Weibull distribution with parameters $\alpha$ and $\beta$, $E(X) = \alpha^{-1/\beta}\,\Gamma\left(\frac{1}{\beta}+1\right)$

$\therefore$ Required mean $= \left(\frac{1}{100}\right)^{-\frac{1}{2}} \Gamma\left(\frac{3}{2}\right)$

$$= 10 \times \frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$= 5\sqrt{\pi}.$$

$$\mathrm{Var}(X) = \alpha^{-\frac{2}{\beta}}\left[\Gamma\left(\frac{2}{\beta}+1\right) - \left\{\Gamma\left(\frac{1}{\beta}+1\right)\right\}^2\right]$$

$$= \left(\frac{1}{100}\right)^{-1}\left[\Gamma(2) - \left\{\Gamma\left(\frac{3}{2}\right)\right\}^2\right]$$

$$= 100\left[1 - \left(\frac{1}{2}\sqrt{\pi}\right)^2\right]$$

$$= 100\left(1 - \frac{\pi}{4}\right)$$

Gamma and Weibull Distribution

GAMMA DISTRIBUTION:

**Definition:** A continuous RV $X$ is said to follow an *Erlang distribution* or *General Gamma distribution* with parameters $\lambda > 0$ and $k > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, & \text{for } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

We note that $\displaystyle\int_0^\infty f(x)\,dx = \frac{\lambda^k}{\Gamma(k)} \int_0^\infty x^{k-1} e^{-\lambda x}\,dx$

$$= \frac{1}{\Gamma(k)} \int_0^\infty t^{k-1} e^{-t}\,dt, \text{ [on putting } \lambda x = t]$$

$$= 1$$

Hence $f(x)$ is a legitimate density function.

## Mean and Variance of Erlang Distribution

The raw moments $\mu_r'$ about the origin of the Erlang distribution are given by

$$\mu_r' = E(X^r)$$

$$= \int_0^\infty \frac{\lambda^k}{\Gamma(k)} x^{k+r-1} e^{-\lambda x}\,dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \cdot \frac{1}{\lambda^{k+r}} \int_0^\infty t^{k+r-1} e^{-t}\,dt, \text{ (on putting } \lambda x = t)$$

$$= \frac{1}{\lambda^r} \frac{\Gamma(k+r)}{\Gamma(k)}$$

$$\therefore \quad \text{Mean} = E(X) = \frac{1}{\lambda} \cdot \frac{\Gamma(k+1)}{\Gamma(k)} = \frac{k}{\lambda}$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

$$= \frac{1}{\lambda^2} \cdot \frac{\Gamma(k+2)}{\Gamma(k)} - \left(\frac{k}{\lambda}\right)^2$$

$$= \frac{1}{\lambda^2} \{k(k+1) - k^2\} = \frac{k}{\lambda^2}$$