ADVANCED ENGINEERING MATHEMATICS

IT (III SEM)

DISTRIBUTIONS

By Dr. Anil Maheshwari

Assistant Professor, Mathematics

Engineering College, Ajmer

BINOMIAL DISTRIBUTION

Binomial distribution is a discrete probability distribution, which is useful, if an experiment is repeated independently on finite occasions e.g.

(i) A coin is tossed 10 times, 15 times,…..

(ii) A die or pair of dice is thrown a number of times etc.

PHYSICAL CONDITIONS FOR BINOMIAL DISTRIBUTION

(*i*) The number of trials 'n' is finite.

(*ii*) Each trial has two exhaustive and mutually disjoint outcomes, called as success and failure.

(*iii*) The trials are independent of each other.

(*iv*) The probability of success is constant for each trial.

For binomial distribution, the probability mass function is given by

$$P(X=r) = {}^{n}C_{r}\, p^{r}\, q^{n-r} \; ; \; r = 0,1,2,\ldots.n,$$

where	$n \rightarrow$	**number of repetitions of an experiment**
	$p \rightarrow$	**probability of success**
	$q \rightarrow$	**probability of failure**
	$r \rightarrow$	**number of successes**

MEAN AND VARIANCE OF BINOMIAL DISTRIBUTION

For the binomial distribution

$$P(r) = {}^nC_r q^{n-r} p^r$$

$$\text{Mean} = E(r) = \mu'_1 = \sum_{r=0}^{n} r.P(r) = \sum_{r=0}^{n} r.\ {}^nC_r\ q^{n-r} P^r$$

$$= 0 + {}^nC_1\ q^{n-1} p + 2\ .\ {}^nC_2\ q^{n-2} p^2 + 3.{}^nC_3\ q^{n-3} p^3 + \ldots + n.\ {}^nC_n\ q^{n-n} p^n$$

$$= nq^{n-1}p + \frac{2n(n-1)}{2.1} q^{n-2} p^2 + \frac{3n(n-1)(n-2)}{3.2.1} q^{n-3} p^3 + \ldots + np^n$$

$$= nq^{n-1}p + n(n-1)q^{n-2}p^2 + \frac{n(n-1)(n-2)}{2.1} q^{n-3} p^3 + \ldots + np^n$$

$$= np\left[q^{n-1} + (n-1)q^{n-2}p + \frac{(n-1)(n-2)}{2.1} q^{n-3} p^2 + \ldots + p^{n-1}\right]$$

$$= np\left[{}^{n-1}C_0 q^{n-1} + {}^{n-1}C_1 q^{n-2} p + {}^{n-1}C_2 q^{n-3} p^2 + \ldots + {}^{n-1}C_{n-1} p^{n-1}\right]$$

$$= np(q+p)^{n-1}$$

$$\Rightarrow \quad \boxed{\text{Mean} = E(r) = \mu'_1 = np}$$

Now, variance $= \sigma^2 = E(r^2) - \{E(r)\}^2 = \mu'_2 - \mu'^2_1 = \sum_{r=0}^{n} r^2 P(r) - \mu'^2_1$

$$\sigma^2 = \sum_{r=0}^{n} [r + r(r-1)]\, P(r) - (np)^2$$

$$= \sum_{r=0}^{n} rP(r) + \sum_{r=0}^{n} r(r-1)\, P(r) - n^2p^2$$

$$\text{Variance} = \sigma^2 = \mu'_1 + \sum_{r=0}^{n} r(r-1)\, P(r) - n^2P^2$$

$$= \mu'_1 + \sum_{r=0}^{n} r(r-1)\, {}^nC_r\, q^{n-r} p^r - n^2p^2$$

$$= \mu'_1 + [2.1\, {}^nC_2\, q^{n-2} p^2 + 3.2\, {}^nC_3\, q^{n-3} p^3 + \ldots + n(n-1)\, {}^nC_n\, p^n] - n^2p^2$$

[Since the contribution due to $r = 0$ and $r = 1$ is zero]

$$= \mu'_1 + \left[2.1\frac{n(n-1)}{2.1}q^{n-2}p^2 + 3.2\frac{n(n-1)(n-2)}{3.2.1}q^{n-3}p^3 + \dots + n(n-1)p^n\right] - n^2p^2$$

$$= \mu'_1 + [n(n-1)q^{n-2}p^2 + n(n-1)(n-2)q^{n-3}p^3 + \dots + n(n-1)p^n] - n^2p^2$$

$$= \mu'_1 + n(n-1)p^2\left[q^{n-2} + (n-2)q^{n-3}p + \dots + p^{n-2}\right] - n^2p^2$$

$$= \mu'_1 + n(n-1)p^2\left[{}^{n-2}C_0q^{n-2} + {}^{n-2}C_1q^{n-3}p + \dots + {}^{n-2}C_{n-2}p^{n-2}\right] - n^2p^2$$

$$= \mu'_1 + n(n-1)p^2(q+p)^{n-2} - n^2p^2$$

$$\text{Variance} = \sigma^2 = np + n(n-1)p^2 - n^2p^2$$

$$= np + n^2p^2 - np^2 - n^2p^2$$

$$= np - np^2$$

$$= np(1-p)$$

$$= npq \qquad [\because 1-p=q]$$

$$\Rightarrow \qquad \boxed{\text{Variance} = \sigma^2 = npq}$$

First four moments about origin

$$\mu_1' = np$$
$$\mu_2' = n(n-1)\,p^2 + np$$
$$\mu_3' = n(n-1)(n-2)\,p^3 + 3n(n-1)\,p^2 + np$$
$$\mu_4' = n(n-1)(n-2)(n-3)\,p^4 + 6n(n-1)(n-2)\,p^3 + 7n(n-1)\,p^2 + np$$

First four central moments

$$\mu_1 = 0$$
$$\mu_2 = npq$$
$$\mu_3 = npq(q-p)$$
$$\mu_4 = npq\left[1+3(n-2)\,pq\right]$$

Karl Pearson's β and γ coefficients

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(1-2p)^2}{npq} \quad ; \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{1+3(n-2)\,pq}{npq}$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{(1-2p)}{\sqrt{npq}} \quad ; \quad \gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} = \frac{1-6pq}{npq}$$

MOMENT GENERATING FUNCTION OF BINOMIAL DISTRIBUTION

$$\boxed{M_x(t) = (q + pe^t)^n}$$

MEAN AND VARIANCE OF BINOMIAL DISTRIBUTION WITH THE HELP OF MGF

MGF of Binomial distribution is given as

$$M_x(t) = (q + pe^t)^n \qquad \text{...(1)}$$

Differentiate equation (1) w.r.t. to 't' we get

$$M'_x(t) = n(q + pe^t)^{n-1} pe^t \qquad \text{...(2)}$$

putting $t = 0$ we get

$$[M'_x(t)]_{t=0} = \mu_1' = n(q + p)^{n-1}p = np \qquad (\because q + p = 1)$$

$$\Rightarrow \qquad \boxed{\text{Mean} = \mu'_1 = np} \qquad \text{...(3)}$$

Again Differentiate equation (2) w.r.t. 't' and putting $t = 0$, we get

$$M''_x(t) = n[(n-1)(q+pe^t)^{n-2}(pe^t)^2 + (q+pe^t)^{n-1}pe^t]$$

$$[M''_x(t)]_{t=0} = n[(n-1)(q+p)^{n-2}p^2 + (q+p)^{n-1}p]$$

$$\Rightarrow \qquad \mu'_2 = n(n-1)p^2 + np \qquad \ldots(4)$$

Now,

$$\text{Variance} = \sigma^2 = \mu'_2 - \mu'^2_1$$

$$= n(n-1)p^2 + np - n^2p^2 \qquad \text{[using (3) and (4)]}$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np(1-p)$$

$$\Rightarrow \qquad \boxed{\text{Variance} = \sigma^2 = npq}$$

Q.1. Ten coins are tossed together. Find the probability of getting at least seven heads.

Sol. If we compare the statements given in the problem with characteristics of binomial distribution, we get $n=10,\ p=\frac{1}{2},\ q=1-p=\frac{1}{2},\ r=7$ or 8 or 9 or 10 i.e. we need to obtain the value of $P(X\geq 7)$

Now, $P(X\geq 7)=P(X=7)+P(X=8)+P(X=9)+P(X=10)$

$$={}^{10}C_7\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^3+{}^{10}C_8\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^2+{}^{10}C_9\left(\frac{1}{2}\right)^9\left(\frac{1}{2}\right)+{}^{10}C_{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^0$$

$$\left[\because P(X=r)={}^nC_r\,p^r\,q^{n-r}\right]$$

$$=\left(\frac{1}{2}\right)^{10}\left({}^{10}C_7+{}^{10}C_8+{}^{10}C_9+{}^{10}C_{10}\right)=\frac{1}{1024}(120+45+10+1)=\frac{11}{64}$$

Q.2. The probability that a bomb dropped from a plane strikes the target is $\frac{1}{5}$. If six bombs are dropped, then what is the probability that at least two bombs will strike the target?

Sol. If we compare the statements given in the problem with characteristics of binomial distribution, we get $n = 6,\ p = \frac{1}{5},\ q = 1 - p = \frac{4}{5},\ r = 2$ or 3 or 4 or 5 or 6 i.e. we need to obtain the value of $P(X \geq 2)$

Now, $P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$

$$= 1 - P(X = 0) - P(X = 1) = 1 - {}^6C_0\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^6 + {}^6C_1\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^5$$

$$\left[\because P(X = r) = {}^nC_r\, p^r\, q^{n-r}\right]$$

$$= 0.34464$$

Q.3. Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six?

Sol. Here, $p =$ the probability of getting 5 or 6 with one dice $= \frac{2}{6} = \frac{1}{3}$ then

$q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$.

Also $n = 6$ and $N = 729$.

(Since dice are in sets of 6 and there are 729 sets.)

The expected number of times that atleast three dice showing 5 or 6 can be obtained by

$$= 729[P(3) + P(4) + P(5) + P(6)]$$

$$= 729\left[{}^6C_3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3 + {}^6C_4\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^4 + {}^6C_5\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^5 + {}^6C_6\left(\frac{1}{3}\right)^6\right]$$

$$= 729\left(\frac{160 + 60 + 12 + 1}{729}\right)$$

$$= 233$$

Q.4. In a precision bombing attack, there is a 50% chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99% chance or better for completely destroying the target?

Sol. As per the given problem, we have $p=\frac{1}{2}, q=1-p=\frac{1}{2}, P(X\geq 2)\geq 0.99$

and we have to find value of n.

As given, $P(X\geq 2)\geq 0.99 \Rightarrow 1-P(X<2)\geq 0.99$

$$\Rightarrow 0.01\geq P(X<2)\Rightarrow 0.01\geq P(X=0)+P(X=1)$$

$$\Rightarrow 0.01\geq {}^{n}C_{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{n}+{}^{n}C_{1}\left(\frac{1}{2}\right)^{1}\left(\frac{1}{2}\right)^{n-1}$$

$$\Rightarrow \frac{1}{100}\geq \frac{1}{2^{n}}(1+n)\Rightarrow \frac{2^{n}}{(1+n)}\geq 100$$

Now, $\frac{2^{10}}{1+10}=\frac{1024}{11}<100$, while $\frac{2^{11}}{1+11}=\frac{2048}{12}>100$

Thus, $n=11$

Q.5. **The mean and variance of binomial distribution are** 4 and $\frac{4}{3}$ **respectively. Find** $P(X \geq 1)$.

Sol. **As given, mean** $= np = 4$ **and variance** $= npq = \frac{4}{3}$

Now, $\frac{npq}{np} = \frac{4/3}{4} \Rightarrow q = \frac{1}{3}$ **and** $p = 1 - q = \frac{2}{3}$

Also, $np = 4 \Rightarrow n = \frac{4}{p} = 6$

So, $P(X \geq 1) = 1 - P(X = 0) = 1 - {}^6C_0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^6$

$\Rightarrow P(X \geq 1) = 1 - \frac{1}{729} = 0.9986$

Q.6. *The following data show the number of seeds germinating out of 10 on damp filter for 80 set of seeds. Fit a Binomial distribution to this data :*

X	:	0	1	2	3	4	5	6	7	8	9	10
f	:	6	20	28	12	8	6	0	0	0	0	0

Sol. Fitting a Binomial Distribution :

X	f	fX
0	6	0
1	20	20
2	28	56
3	12	36
4	8	32
5	6	30
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
	$N = \Sigma f = 80$	$\Sigma fX = 174$

Here $\overline{X} = \frac{\Sigma fX}{N} = \frac{174}{80}$

or $\overline{X} = 2.175$

$$\text{Mean} = np = 2.175$$

$$p = \frac{2.175}{10} = 0.2175$$

Then, $q = 1 - p = 1 - .2175 = .7825$

The theoretical frequencies are given below :

X	Theoretical Frequencies $N \times {}^nC_r q^{n-r} p^r$		
0	$80 \times {}^{10}C_0 (.7825)^{10} (.2175)^0$	=	6.9
1	$80 \times {}^{10}C_1 (.7825)^9 (.2175)^1$	=	19.1
2	$80 \times {}^{10}C_2 (.7825)^8 (.2175)^2$	=	24.0
3	$80 \times {}^{10}C_3 (.7825)^7 (.2175)^3$	=	17.8
4	$80 \times {}^{10}C_4 (.7825)^6 (.2175)^4$	=	8.6
5	$80 \times {}^{10}C_5 (.7825)^5 (.2175)^5$	=	2.9
6	$80 \times {}^{10}C_6 (.7825)^4 (.2175)^6$	=	0.7
7	$80 \times {}^{10}C_7 (.7825)^3 (.2175)^7$	=	0.1
8	$80 \times {}^{10}C_8 (.7825)^2 (.2175)^8$	=	0.0
9	$80 \times {}^{10}C_9 (.7825)^1 (.2175)^{10}$	=	0.0

POISSON DISTRIBUTION

Poisson distribution is a discrete probability distribution.

Following are the situations where Poisson distribution may be successfully employed :

(*a*) Number of cars passing a trisection per minute during the busy hours of a day.

(*b*) Number of air accidents in some unit of time in a country.

(*c*) Number of suicides reported in a particular city in India.

(*d*) Number of wrong telephone calls received by a office in a day *etc.*

Poisson distribution is a limiting case of the Binomial distribution under the following conditions :

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

(*ii*) p, the probability of success for each trial is indefinitely small, *i.e.*, $p \to 0$.

(*iii*) $np = \lambda$, (say) is a finite positive real number.

For Poisson distribution, the probability mass function is given by

$$P(X = x) = \frac{e^{-\lambda} \lambda^{x}}{\lfloor x} \quad ; \quad x = 0, 1, 2 \ldots ; \lambda > 0$$

where $\lambda \rightarrow$ parameter of the distribution

$e \rightarrow$ a constant, whose value is approximately equal to 2.7183

$x \rightarrow$ number of successes

MEAN AND VARIANCE OF THE POISSON DISTRIBUTION

For the Poisson distribution,

$$P(x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$\text{Mean} = E(x) = \mu'_1 = \sum_{x=0}^{\infty} xP(x) = \sum_{x=0}^{\infty} x . \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\sum_{x=1}^{\infty} \frac{\lambda^x}{\lfloor x-1} = e^{-\lambda}\left(\lambda + \frac{\lambda^2}{\lfloor 1} + \frac{\lambda^3}{\lfloor 2} + \ldots\right)$$

$$= \lambda e^{-\lambda}\left(1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \ldots\right) = \lambda e^{-\lambda} e^{\lambda}$$

$$\Rightarrow \qquad \boxed{\text{Mean} = E(x) = \mu'_1 = \lambda} \qquad \ldots(1)$$

Now, $$\text{Variance} = \sigma^2 = E(x^2) - \{E(x)\}^2$$
$$\sigma^2 = \mu'_2 - {\mu'_1}^2 \qquad \text{...(2)}$$

Now, $$\mu'_2 = E(x^2) = \sum_{x=0}^{\infty} x^2 P(x)$$

$$= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\lambda}\lambda^x}{\lfloor x} = \sum_{x=0}^{\infty} [x(x-1)+x] \cdot \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{\lfloor x} + \sum_{x=0}^{\infty} \frac{xe^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda} \sum_{x=2}^{\infty} \frac{x(x-1)\lambda^x}{x(x-1)\lfloor x-2} + \lambda = e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{\lfloor x-2} + \lambda$$

$$= e^{-\lambda}\lambda^2 \left(1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \frac{\lambda^3}{\lfloor 3} + \dots\right) + \lambda = e^{-\lambda}\lambda^2 \cdot e^{\lambda} + \lambda$$

$$\Rightarrow \quad \mu'_2 = \lambda^2 + \lambda \qquad \text{...(3)}$$

Now, from equation (2)

$$\sigma^2 = \mu'_2 - {\mu'_1}^2$$
$$= \lambda^2 + \lambda - \lambda^2 \qquad \text{[using (1) and (3)]}$$

$$\Rightarrow \quad \boxed{\text{Variance} = \sigma^2 = \lambda}$$

First four moments about origin

$$\mu_1' = \lambda$$
$$\mu_2' = \lambda^2 + \lambda$$
$$\mu_3' = \lambda^3 + 3\lambda^2 + \lambda$$
$$\mu_4' = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

First four central moments

$$\mu_1 = 0$$
$$\mu_2 = \lambda$$
$$\mu_3 = \lambda$$
$$\mu_4 = 3\lambda^2 + \lambda$$

Karl Pearson's β and γ coefficients

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \quad ; \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2 + \lambda}{\lambda^2} = 3 + \frac{1}{\lambda}$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}} \quad ; \quad \gamma_2 = \beta_2 - 3 = 3 + \frac{1}{\lambda} - 3 = \frac{1}{\lambda}$$

MOMENT GENERATING FUNCTION OF POISSON DISTRIBUTION

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

MEAN AND VARIANCE OF POISSON DISTRIBUTION WITH THE HELP OF MGF

MGF of Poisson distribution is given as

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

Differentiate equation (1) w.r.t. to 't' we get

$$M'_x(t) = e^{\lambda(e^t - 1)}\lambda e^t \quad \ldots(2)$$

putting $t = 0$ we get

$$[M'_x(t)]_{t=0} = \mu_1' = \lambda$$

$$\Rightarrow \quad \boxed{\text{Mean} = \mu'_1 = \lambda} \quad \ldots(3)$$

Again Differentiate equation (2) w.r.t. 't' and putting $t = 0$, we get

$$M''_x(t) = e^{\lambda(e^t - 1)} \lambda e^t + e^{\lambda(e^t - 1)} \lambda^2 e^{2t}$$

$$[M''_x(t)]_{t=0} = \lambda + \lambda^2$$

$$\Rightarrow \qquad \mu'_2 = \lambda + \lambda^2 \qquad \text{...(4)}$$

Now,

$$\text{Variance} = \sigma^2 = \mu'_2 - \mu'^2_1$$

$$= \lambda + \lambda^2 - \lambda^2 \qquad \text{[using (3) and (4)]}$$

$$= \lambda$$

$$\Rightarrow \qquad \boxed{\text{Variance} = \sigma^2 = \lambda}$$

Q.1. A factory produces razor blades. The probability of its being defective is $1/500$. In 10000 packets of 10 blades each. Use Poisson distribution to calculate the approximate number of packets (a) having no defective blade

(b) having one defective blade $\left(e^{-0.02} = 0.9802\right)$

Sol. As per the given problem, we have

$$\lambda = np = 10\left(\frac{1}{500}\right) = 0.02 \Rightarrow e^{-\lambda} = e^{-0.02} = 0.9802$$

Now, required probabilities are given as

$$P(X=0) = \frac{e^{-\lambda}\lambda^0}{\lfloor 0} = 0.9802 \text{ and } P(X=1) = \frac{e^{-\lambda}\lambda^1}{\lfloor 1} = 0.0196$$

Thus, the requires solutions are given as

(a) $10000 \times P(X=0) = 10000 \times 0.9802 = 9802$

(b) $10000 \times P(X=1) = 10000 \times 0.0196 = 196$

Q.2. Assume that the probability of an individual coalminer being killed in a mine accident during a year is $1/2400$. Use Poisson distribution to calculate the probability that in a mine employing 200 miners, there will be at least one fatal accident in a year.

Sol. As per the given problem, we have

$$\lambda = np = 200\left(\frac{1}{2400}\right) = \frac{1}{12} \Rightarrow e^{-\lambda} = e^{-1/12} = 0.9201$$

Now, required probability is given as

$$P(X \geq 1) = P(X=1) + P(X=2) + \ldots.. + P(X=200)$$

$$1 - P(X=0) = 1 - \frac{e^{-\lambda}\lambda^0}{\underline{|0}} = 1 - 0.9201 = 0.0799$$

Q.3. **A car hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as Poisson distribution with mean $^3/_2$. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused.** $\left(e^{-1.5} = 0.2231\right)$

Sol. **As per the given problem, we have mean of Poisson distribution** $= {}^3/_2 \Rightarrow \lambda = 1.5$

$$\Rightarrow e^{-\lambda} = e^{-1.5} = 0.2231$$

Now, proportion of days on which neither car is used

$$= P(X=0) = \frac{e^{-\lambda}\lambda^0}{\lfloor 0} = 0.2231$$

and proportion of days on which some demand is refused

$$= P(X \geq 3) = 1 - P(X=0) - P(X=1) - P(X=2)$$

$$= 1 - \frac{e^{-\lambda}\lambda^0}{\lfloor 0} - \frac{e^{-\lambda}\lambda^1}{\lfloor 1} - \frac{e^{-\lambda}\lambda^2}{\lfloor 2}$$

$$= 1 - 0.2231 - 0.3347 - 0.2510 = 0.1912$$

Q.4. **If** X **is a Poisson variable such that** $P(X=2)=9P(X=4)+90P(X=6)$. **Find the standard deviation.**

Sol. **As given,** $P(X=2)=9P(X=4)+90P(X=6)$

$$\Rightarrow \frac{e^{-\lambda}\lambda^2}{\lfloor 2}=9\frac{e^{-\lambda}\lambda^4}{\lfloor 4}+90\frac{e^{-\lambda}\lambda^6}{\lfloor 6}$$

$$\Rightarrow \frac{\lambda^2}{2}=9\frac{\lambda^4}{24}+90\frac{\lambda^6}{720} \Rightarrow 1=3\frac{\lambda^2}{4}+\frac{\lambda^4}{4}$$

$$\Rightarrow \lambda^4+3\lambda^2-4=0 \Rightarrow \lambda^2=1 \quad \left(\because \lambda^2 \neq -4\right)$$

$$\Rightarrow \text{variance}=\lambda=1 \quad (\because \text{variance can't be negative})$$

So, standard deviation $=\sqrt{\text{variance}}=1$

Q.5. *Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 200 army crops. The distribution of deaths was as follows:*

No. of deaths	:	*0*	*1*	*2*	*3*	*4*	*Total*
Frequency	:	*109*	*65*	*22*	*3*	*1*	*200*

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

Sol. **Calculation for Mean :**

X	f	fX
0	109	0
1	65	65
2	22	44
3	3	9
4	1	4
	$\Sigma f = 200$	$\Sigma fX = 122$

Here $$\overline{X} = \frac{\Sigma fX}{\Sigma f} = \frac{122}{200} = 0.61 = \lambda$$

This is the parameter (λ) of the Poisson distribution.

Calculation for expected frequencies

X	$P(X = x) = \dfrac{e^{-\lambda}\,\lambda^x}{\lfloor x}$	Frequency $NP(x)$
0	$\dfrac{e^{-0.61}\,(.61)^0}{\lfloor 0} = .5432$	$200 \times .5432 = 108.64$
1	$\dfrac{e^{-0.61}\,(.61)^1}{\lfloor 1} = .3313$	$200 \times .3313 = 66.27$
2	$\dfrac{e^{-0.61}\,(.61)^2}{\lfloor 2} = .101$	$200 \times .101 = 20.21$
3	$\dfrac{e^{-0.61}\,(.61)^3}{\lfloor 3} = .021$	$200 \times .021 = 4.11$
4	$\dfrac{e^{-0.61}\,(.61)^4}{\lfloor 4} = .003$	$200 \times .003 = .63$

Hence theoretical frequencies are 109, 66, 20, 4, 1

NORMAL DISTRIBUTION

Normal distribution is the most important continuous probability distribution used in Statistics.

The normal distribution has wide applications in the theory of Statistics. The tests of significance for large samples are based on this distribution.

Normal distribution is a limiting case of the binomial distribution, when n is very large and p and q are not very small.

Normal distribution can also be obtained as a limiting case of Poission distribution with the parameter $\lambda \to \infty$.

For normal distribution, the probability density function is given by

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

$$-\infty < \mu < \infty, \sigma > 0$$

with parameters μ (mean) and σ^2 (variance).

Notations : $X \sim N(\mu, \sigma)$ or $X \sim N(\mu, \sigma^2)$

i.e. X is normally distributed with parameter μ and σ

STANDARD NORMAL DISTRIBUTION

If $X \sim N(\mu, \sigma)$, then

$$Z = \frac{X - \mu}{\sigma}$$ is called standard normal variate.

i.e.

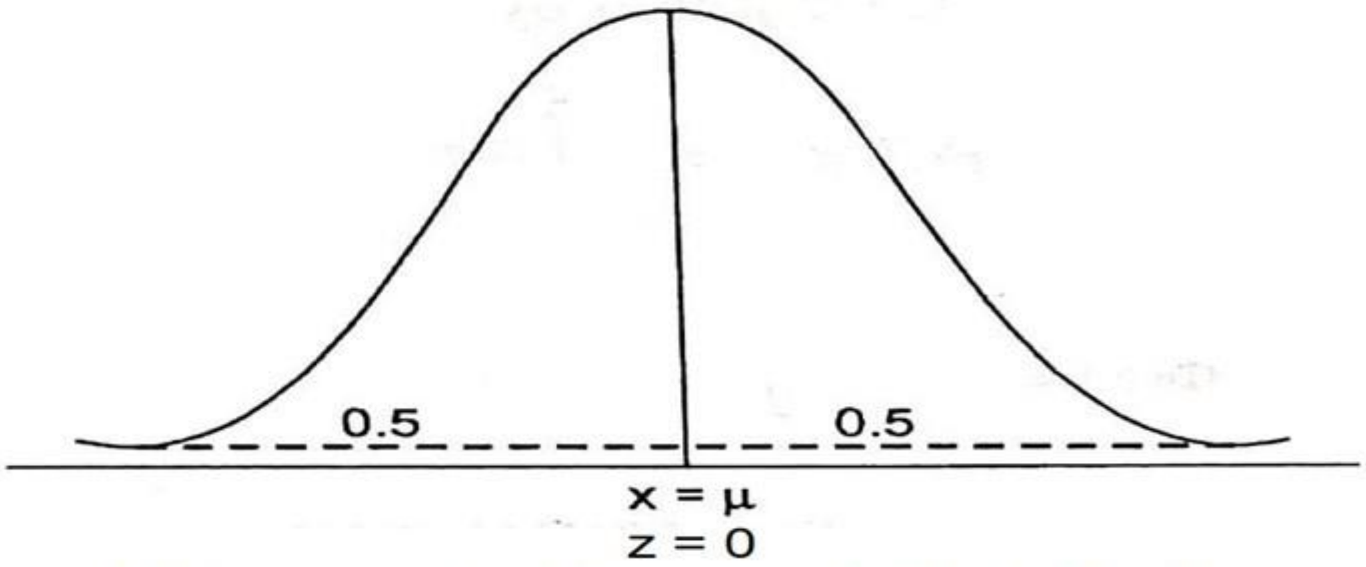
$$\boxed{Z = \frac{X - \text{mean}}{\sqrt{\text{Var.}}}}$$

The notation of standard normal variate is

$$Z \sim N(0, 1)$$

i.e. mean $(Z) = 0$ and variance $(Z) = 1$

Graph of Normal Distribution



Characteristics of the graph of normal distribution

(*i*) The curve is bell shapped and symmetrical about the line $x = \mu$.

(*ii*) Since $f(x)$ being the probability, can never be negative. Hence no portion of the curve lies below the x-axis.

(*iii*) The ordinate at the mean of the distribution divides the total area under the normal curve into two equal parts. Since, total area under normal probability curve is 1, the area to the right of the ordinate as well as to the left of the ordinate is 0.5.

MEAN AND VARIANCE OF NORMAL DISTRIBUTION

Let $X \sim N(\mu, \sigma^2)$, then

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

$$\therefore \quad \text{Mean} = E(X) = \mu'_1 = \int_{-\infty}^{\infty} x f(x)\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad ; -\infty < x < \infty$$

Let $\quad z = \frac{x-\mu}{\sigma} \Rightarrow x = \mu + \sigma z \Rightarrow dx = \sigma\, dz$

Hence, $\quad \text{mean} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma z) e^{-z^2/2}\, dz$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu\, e^{-z^2/2}\, dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\, e^{-z^2/2}\, dz$$

$$= \frac{2\mu}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-z^2/2}\, dz + 0 \qquad \left[\begin{array}{l}\text{Since } ze^{-z^2/2} \text{ is an odd function of } z \\ \text{and } \quad e^{-z^2/2} \text{ is an even function of } z\end{array}\right]$$

$$= \frac{2\mu}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-z^2/2}\, dz$$

$$= \frac{2\mu}{\sqrt{2\pi}} \int_0^\infty e^{-t} \frac{dt}{\sqrt{2t}} \qquad \left[\text{Let } \frac{z^2}{2} = t \Rightarrow z\,dz = dt \Rightarrow dz = \frac{dt}{\sqrt{2t}} \right]$$

$$= \frac{\mu}{\sqrt{\pi}} \int_0^\infty e^{-t} t^{-1/2} dt \qquad = \frac{\mu}{\sqrt{\pi}} \int_0^\infty e^{-t} t^{1/2-1} dt$$

$$= \frac{\mu}{\sqrt{\pi}} \Gamma\frac{1}{2} \qquad \left[\textbf{Since } \int_0^\infty e^{-x} x^{n-1} dx = \Gamma n \right]$$

$$= \frac{\mu}{\sqrt{\pi}} \times \sqrt{\pi} = \mu$$

$$\Rightarrow \qquad \boxed{\text{Mean} = E(X) = \mu'_1 = \mu}$$

Now, Variance = $E[X - \text{mean}]^2 = E[X - \mu]^2$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Let $z = \dfrac{x-\mu}{\sigma} \Rightarrow x = \mu + \sigma z \Rightarrow dx = \sigma\, dz$

Hence, $\text{variance} = \dfrac{1}{\sigma\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} \sigma^2 z^2 e^{-z^2/2}\, \sigma\, dz = \dfrac{\sigma^2}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} z^2 e^{-z^2/2}\, dz$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} z^2 e^{-z^2/2}\, dz$$

[Since $z^2 e^{-z^2/2}$ is an even function of z]

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^\infty 2t\, e^{-t} \cdot \frac{dt}{\sqrt{2t}} \quad \left[\text{Let } \frac{z^2}{2} = t \Rightarrow z\,dz = dt \Rightarrow dz = \frac{dt}{\sqrt{2t}} \right]$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^\infty e^{-t} . (2t)^{1/2} dt = \frac{2\sigma^2}{\sqrt{\pi}} \int_0^\infty e^{-t} . t^{1/2} dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_0^\infty e^{-t} . t^{\frac{3}{2}-1}\, dt \quad \left[\text{Since } \int_0^\infty e^{-x} x^{n-1} dx = \Gamma n \right]$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \Gamma \frac{3}{2} = \frac{2\sigma^2}{\sqrt{\pi}} . \frac{1}{2} \Gamma \frac{1}{2} = \frac{\sigma^2}{\sqrt{\pi}} \times \sqrt{\pi} = \sigma^2$$

$$\Rightarrow \quad \boxed{\text{Variance} = \sigma^2}$$

Hence, Standard Deviation $= \sqrt{\text{Variance}} = \sigma$

Odd order central moments

$$\mu_{2n+1} = 0 \quad ; \quad n = 0, 1, 2, \ldots.$$

i.e. all central moments of odd order become zero.

Even order central moments

$$\mu_{2n} = 1.3.5\ldots.(2n-1)\sigma^{2n} \quad ; \quad n = 1, 2, 3, \ldots.$$

First four central moments

$$\mu_1 = 0 \qquad \mu_2 = \sigma^2$$

$$\mu_3 = 0 \qquad \mu_4 = 3\sigma^4$$

Karl Pearson's β and γ coefficients

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{\sigma^6} = 0 \quad ; \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\sigma^4}{\sigma^4} = 3$$

$$\gamma_1 = \sqrt{\beta_1} = 0 \quad ; \quad \gamma_2 = \beta_2 - 3 = 3 - 3 = 0$$

MOMENT GENERATING FUNCTION OF NORMAL DISTRIBUTION

$$\boxed{M_x(t) = e^{\left(t\mu + \frac{1}{2}t^2\sigma^2\right)}}$$

MEAN AND VARIANCE OF NORMAL DISTRIBUTION WITH THE HELP OF MGF

MGF of normal distribution is given as

$$M_x(t) = e^{\left(t\mu + \frac{1}{2}t^2\sigma^2\right)} \qquad \ldots(1)$$

Differentiate equation (1) w.r.t. to 't' we get

$$M_X'(t) = e^{\left(t\mu + \frac{1}{2}t^2\sigma^2\right)}(\mu + t\sigma^2) \qquad \ldots(2)$$

putting $t = 0$ we get

$$[M_X'(t)]_{t=0} = \mu_1' = \mu$$

$$\Rightarrow \qquad \boxed{\text{Mean} = \mu'_1 = \mu} \qquad \ldots(3)$$

Again Differentiate equation (2) w.r.t. 't' and putting $t = 0$, we get

$$M''_x(t) = e^{\left(t\mu + \frac{1}{2}t^2\sigma^2\right)}(\mu + t\sigma^2)^2 + e^{\left(t\mu + \frac{1}{2}t^2\sigma^2\right)} \times (\sigma^2)$$

$$[M''_x(t)]_{t=0} = \mu^2 + \sigma^2$$

$$\Rightarrow \quad \mu'_2 = \mu^2 + \sigma^2 \qquad \ldots(4)$$

Now, $\quad$ Variance $= \mu'_2 - \mu'^2_1$

$$= \mu^2 + \sigma^2 - \mu^2 \qquad \text{[using (3) and (4)]}$$

$$= \sigma^2$$

$$\Rightarrow \quad \boxed{\text{Variance} = \sigma^2}$$

Procedure for solving problems based over "normal distribution"

Step – 1 Convert given problem of the form $P(x_1 \le X \le x_2)$ into

The form $P(z_1 \le z \le z_2)$ using $z = \dfrac{X-\mu}{\sigma}$ with given

values of μ and σ. Similarly problems of type $P(x_1 \le X)$

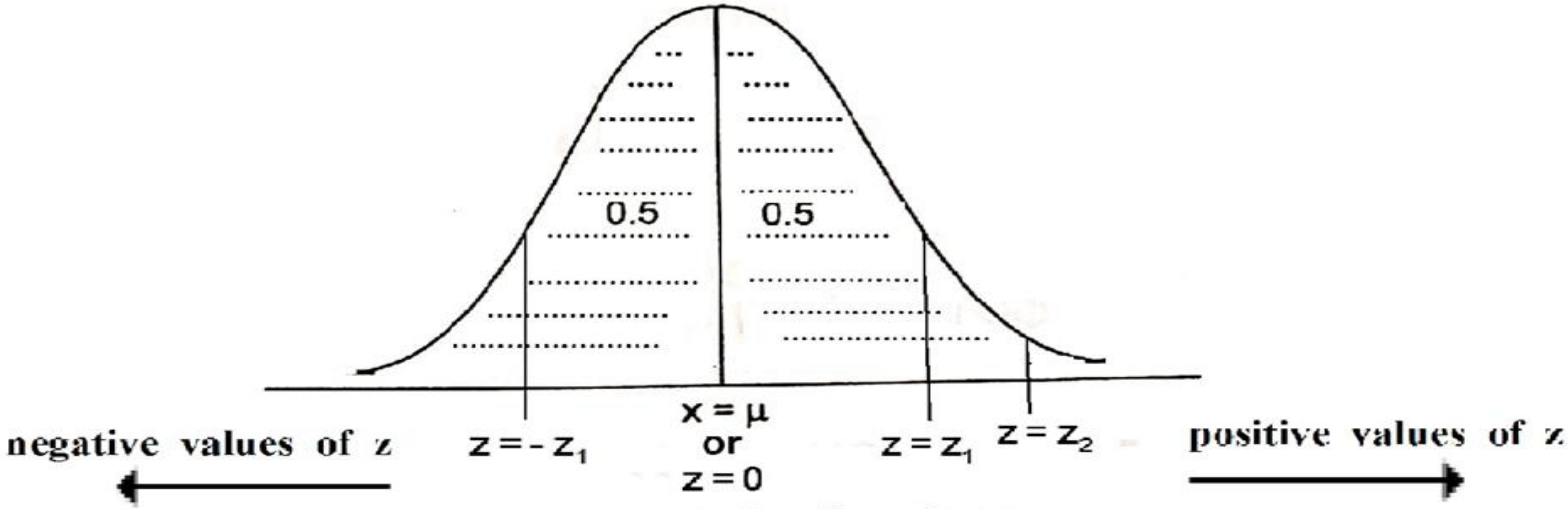
or $P(X \le x_2)$ can be converted into $P(z_1 \le z)$ or

$P(z \le z_2)$ respectively.

Step – 2

Use the following curve of normal variable's p.d.f.

$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$ **to convert all probabilities in** $P(0\le z\le a)$; $a\in R^+$ **form.**



Normal Distribution Curve

Now, we discuss several cases to understand the step - 2 using the normal distribution curve as given below:

Case - 1 $P(-z_1 \le z \le 0) = P(0 \le z \le z_1)$

Case - 2 $P(-z_1 \le z \le z_2) = P(-z_1 \le z \le 0) + P(0 \le z \le z_2)$

$$= P(0 \le z \le z_1) + P(0 \le z \le z_2)$$

Case - 3 $P(-z_1 \le z \le z_1) = 2P(0 \le z \le z_1)$

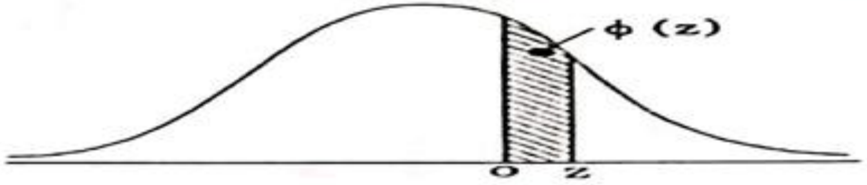
Case - 4 $P(z_1 \le z \le z_2) = P(0 \le z \le z_2) - P(0 \le z \le z_1)$

Case - 5 $P(z \ge z_2) = 0.5 - P(0 \le z \le z_2)$

Case - 6 $P(z \le -z_1) = P(z \ge z_1) = 0.5 - P(0 \le z \le z_1)$

Step - 3

Table: Area Under Standard Normal Curve



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1218	0.1255	0.1293	0.1331	0.1368	0.1496	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1722	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2703	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3181	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4302	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4405	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4852	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4926	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

e.g. $P(0 < z < 1.18) = 0.3810$

Q.1. **If X is a normal variable with mean 30 and standard deviation 5. Find the probabilities**

(a) $26 \le X \le 40$ **(b)** $X \ge 45$ **(c)** $|X - 30| > 5$

Sol. **As per the given problem, we have** $\mu = 30$ and $\sigma = 5$.

So, $z = \dfrac{X - \mu}{\sigma} \Rightarrow X = 30 + 5z$

(a) $P(26 \le X \le 40) = P(26 \le 30 + 5z \le 40) = P(-0.8 \le z \le 2)$

$= P(-0.8 \le z \le 0) + P(0 \le z \le 2)$

$= P(0 \le z \le 0.8) + P(0 \le z \le 2)$

$= 0.2881 + 0.4772 = 0.7653$ [using normal table]

(b) $P(X \ge 45) = P(30 + 5z \ge 45) = P(z \ge 3) = 0.5 - P(0 \le z \le 3)$

$= 0.5 - 0.4987 = 0.0013$ [using normal table]

(c) $P(|X - 30| > 5) = P(|30 + 5z - 30| > 5) = P(|z| > 1)$

$= P(z < -1) + P(z > 1) = P(z > 1) + P(z > 1) = 2P(z > 1)$

$= 2[0.5 - P(0 < z < 1)] = 2(0.5 - 0.3413) = 0.3174$

[using normal table]

Q.2. *Assume the mean height of soldiers to be 68.22 inches with a variance of 10.8 inches. How many soldiers in a regiment of 1000 would you expect to be over six feet tall ?*

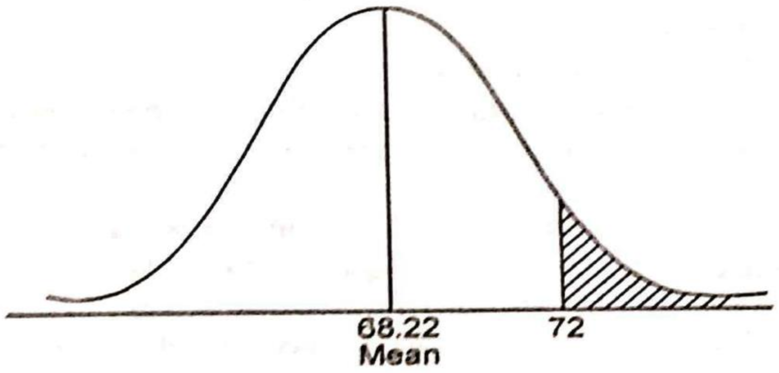
Sol. Given $\mu = 68.22$ inches, $\sigma = \sqrt{10.8} = 3.286$ inches.

Standard normal variate $z = \dfrac{X - \mu}{\sigma}$

Here, $X = 72$ inches

So, $z = \dfrac{72 - 68.22}{3.286} = 1.15$

$$P(X > 72) = P(z > 1.15)$$
$$= 0.5 - P(0 < z < 1.15)$$
$$= 0.5 - 0.3749 \quad \text{(using normal table)}$$
$$= 0.1251$$



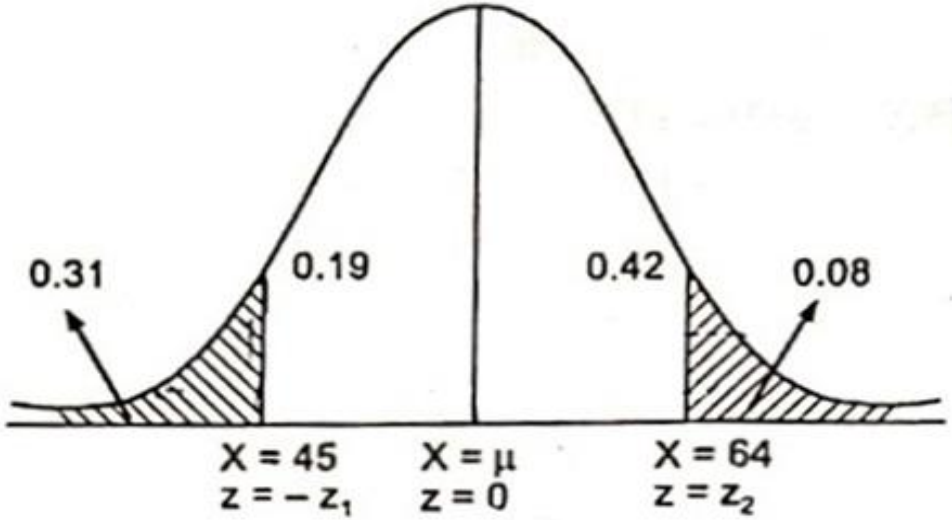
Hence the probability of getting soldiers above six feet is 0.1251 and out of 1000 soldiers the expectation is $1000 \times 0.1251 = 125.1$ or 125.

So, expected number of soldiers over six feet tall is 125.

Q.3. *In a normal distribution, 31% items are under 45 and 8% are over 64. What is the mean and standard deviation of the distribution?*

Sol. Let μ and σ be the mean and standard deviation of the distribution.

Given that $P(X < 45) = 0.31$ and $P(X > 64) = 0.08$



$P(X < 45) = 0.31$ indicates that $X = 45$ should be on the left side of mean μ.

So, that $$-z_1 = \frac{45-\mu}{\sigma} \qquad ...(1)$$

Again $P(X > 64) = 0.08$ indicates that $X = 64$ should be on the right side of mean μ.

So, that $$z_2 = \frac{64-\mu}{\sigma} \qquad ...(2)$$

$$P(X < 45) = 0.31$$

$\Rightarrow$ $$P(z < -z_1) = 0.31$$

or $$0.5 - P(-z_1 < z < 0) = 0.31$$

or $$P(-z_1 < z < 0) = 0.19$$

or $$P(0 < z < z_1) = 0.19$$ [since $P(0 < z < z_1) = P(-z_1 < z < 0)$]

$\Rightarrow$ $z_1 = 0.50$ (From the area under normal curve table)

$$P(X > 64) = 0.08$$

$\Rightarrow$ $$P(z > z_2) = 0.08$$

or $$0.5 - P(0 < z < z_2) = 0.08$$

or $$P(0 < z < z_2) = 0.42$$

$\Rightarrow$ $z_2 = 1.41$ (From the area under normal curve table)

equations (1) and (2) become

$\Rightarrow$ $$\frac{45 - \mu}{\sigma} = -0.50 \qquad ...(3)$$

and $$\frac{64 - \mu}{\sigma} = 1.41 \qquad ...(4)$$

Dividing these two equations, we get

$$-\frac{1.41}{0.50} = \frac{64 - \mu}{45 - \mu}$$

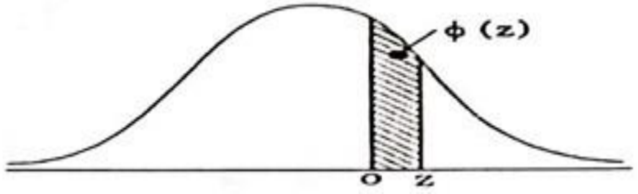
or $$1.91\,\mu = 95.45$$

or $$\mu \cong 50$$

on putting the value of μ into equation (3), we have $\frac{45 - 50}{\sigma} = -0.50$

or $$\sigma = \frac{5}{0.50} = 10$$

Table: Area Under Standard Normal Curve



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1218	0.1255	0.1293	0.1331	0.1368	0.1496	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1722	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2703	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3181	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4302	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4405	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4852	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4926	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

UNIFORM DISTRIBUTION

Uniform distribution is a continuous probability distribution. As in this distribution, random variable can be assigned only uniform (constant) values, it is called uniform distribution.

It is also known as rectangular distribution, because its curve describes a rectangle between the x - axis and the ordinates at points $x = a$ and $x = b$.

Uniform distribution is mainly useful, when the observed values in a set of data are equally sprade across the range of the data set. For example, drawind a card from a deck of cards, tossing a coin, a passanger waitinig for a bus or a train in a system, where buses or trains arrive after equal interval etc.

For uniform distribution, the probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & ; \quad a < x < b \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Notations : $X \sim U(a,b)$ or $X \sim R(a,b)$

i.e. X is uniformally distributed on the interval (a,b)

MEAN AND VARIANCE OF UNIFORM DISTRIBUTION

$$\text{Mean} = E(X) = \mu'_1 = \int_a^b x\, f(x)\, dx = \int_a^b x \frac{1}{b-a} dx$$

$$= \frac{1}{b-a}\int_a^b x\, dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{2(b-a)}(b^2 - a^2) = \frac{1}{2(b-a)}(b+a)(b-a)$$

$$\Rightarrow \quad \boxed{\text{Mean} = E(X) = \mu'_1 = \frac{b+a}{2}}$$

Now, $$\mu'_2 = \int_a^b x^2 f(x)\,dx = \frac{1}{b-a}\int_a^b x^2 dx$$

$$= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{1}{3(b-a)}(b^3 - a^3)$$

$$= \frac{(b-a)(b^2+ba+a^2)}{3(b-a)}$$

$$\Rightarrow \qquad \mu'_2 = \frac{1}{3}(b^2+ba+a^2)$$

Hence, $$\text{variance} = \sigma^2 = \mu'_2 - \mu'^2_1$$

$$= \frac{1}{3}(b^2+ba+a^2) - \frac{1}{4}(b+a)^2$$

$$= \frac{1}{12}[4(b^2+ba+a^2) - 3(b+a)^2]$$

$$= \frac{1}{12}(b^2 - 2ba + a^2)$$

$$\Rightarrow \qquad \boxed{\text{variance} = \sigma^2 = \frac{(b-a)^2}{12}}$$

MOMENT GENERATING FUNCTION OF UNIFORM DISTRIBUTION

$$\boxed{M_X(t) = \frac{1}{t(b-a)}\left(e^{tb} - e^{ta}\right)}$$

MEAN AND VARIANCE OF UNIFORM DISTRIBUTION WITH THE HELP OF MGF

MGF of uniform distribution is given as

$$M_X(t) = \frac{1}{t(b-a)}\left(e^{tb} - e^{ta}\right)$$

$$= \frac{1}{t(b-a)}\left[\left(1 + \frac{tb}{\lfloor 1} + \frac{t^2 b^2}{\lfloor 2} + \ldots\right) - \left(1 + \frac{ta}{\lfloor 1} + \frac{t^2 a^2}{\lfloor 2} + \ldots\right)\right]$$

$$= \frac{1}{t(b-a)}\left[t(b-a) + \frac{t^2}{2}\left(b^2 - a^2\right) + \frac{t^3}{6}\left(b^3 - a^3\right) + \ldots\right]$$

$$= 1 + \frac{t}{2}(b+a) + \frac{t^2}{6}\left(b^2 + ba + a^2\right) + \ldots..$$

$$\Rightarrow \quad \text{coefficient of } \frac{t}{\lfloor 1} = \mu_1' = \frac{1}{2}(b+a)$$

and $\quad$ coefficient of $\dfrac{t^2}{\lfloor 2} = \mu_2' = \dfrac{1}{3}\left(b^2 + ba + a^2\right)$

$$\Rightarrow \quad \boxed{\text{mean } = \mu_1' = \frac{1}{2}(b+a)}$$

and $\quad$ variance $= \mu_2' - \mu_1'^2$

$$\Rightarrow \quad \boxed{\text{variance } = \frac{1}{12}(b-a)^2}$$

Q. 1. *If a random variable λ is uniformly distributed over (0, 10), then find the probability that the roots of the equation $4x^2 + 4\lambda x + (\lambda + 2) = 0$ are real.*

Sol. Since λ is uniformly distributed over (0, 10) so its probability density function is given by :

$$f(\lambda) = \begin{cases} \dfrac{1}{10}, & 0 \le \lambda \le 10 \\ 0, & \text{otherwise} \end{cases}$$

The roots of the equation $4x^2 + 4\lambda x + (\lambda + 2) = 0$ are real if its discriminant $(4\lambda)^2 - 4\,(4)\,(\lambda + 2)$ is non-negative.

$$16\,\lambda^2 - 16\,(\lambda + 2) \ge 0$$

or $$\lambda^2 - \lambda - 2 \ge 0$$

or $$(\lambda - 2)\,(\lambda + 1) \ge 0$$

Required probability $= P\,[(\lambda - 2)(\lambda + 1) \geq 0]$

$= P\,[\{\lambda - 2 \geq 0 \text{ and } \lambda + 1 \geq 0\} \text{ or } \{\lambda - 2 \leq 0 \text{ and } \lambda + 1 \leq 0\}]$

$$= P\,[\{\lambda \geq 2 \text{ and } \lambda \geq -1\} \text{ or } \{\lambda \leq 2 \text{ and } \lambda \leq -1\}]$$

$$= P\,(\lambda \geq 2) \text{ or } (\lambda \leq -1)$$

$$= P\,(\lambda \geq 2) + P\,(\lambda \leq -1)$$

$$= \int_{2}^{10} f(\lambda)\, d\lambda + \int_{-\infty}^{-1} f(\lambda)\, d\lambda$$

$$= \int_{2}^{10} \frac{1}{10}\, d\lambda + \int_{-\infty}^{-1} 0\, d\lambda = \frac{1}{10}.8 = \frac{4}{5}$$

Q. 2. *Buses arrive at a particular bus stop after every 15 minutes, starting from 6 A.M. If a passanger arrives at the stop at a random time that is uniformly distributed between 9 to 9 : 30 A.M., then find the probability that he waits for :*

(a) less than 5 minutes for a bus and

(b) atleast 10 minutes for a bus.

Sol. Let X denote the time in minutes of arrival of a person past 9 A.M.

X is uniformly distributed in the interval $(0, 30)$.

Probability density function of X is given by

$$f(x) = \begin{cases} \dfrac{1}{30}, & 0 \le x \le 30 \\ 0, & \text{otherwise} \end{cases}$$

Buses arrive at a particular bus stop after every 15 minutes, starting *from 6 A.M*; hence after 9 : 00 A.M there will be two buses, one at 9 : 15 A.M and other at 9 : 30 A.M.

(a) A person will have to wait for less than 5 minutes if he arrives between 9 : 10 A.M to 9 : 15 A.M or 9 : 25 A.M to 9 : 30 A.M.

Hence, required probability

$$= P(10 < X < 15) + P(25 < X < 30)$$

$$= \int_{10}^{15} f(x)dx + \int_{25}^{30} f(x)dx = \frac{1}{30}\int_{10}^{15} dx + \frac{1}{30}\int_{25}^{30} dx$$

$$= \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

(b) The person will have to wait atleast 10 minutes if he arrives at the stop between 9 : 00 A.M to 9 : 05 A.M and 9 : 15 A.M to 9 : 20 A.M.

Hence, required probability

$$= P(0 < X < 5) + P(15 < X < 20)$$

$$= \int_{0}^{5} f(x)\,dx + \int_{15}^{20} f(x)dx = \frac{1}{30}\int_{0}^{5} dx + \frac{1}{30}\int_{0}^{5} dx$$

$$= \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

EXPONENTIAL DISTRIBUTION

Exponential distribution is a continuous probability distribution

It is often used in the cases, where the time elapsed between events is considered. For example, the length of long distance business telephone calls, the time calculation for a possible earthquake or a possible cyclone etc.

For exponential distribution, the probability density function is given by

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

Here, the parameter λ is strictly positive i.e. $\lambda > 0$.

The corresponding cumulative distribution function $F(x)$ is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & ; x \geq 0 \\ 0 & ; x < 0 \end{cases}$$

MEAN AND VARIANCE OF EXPONENTIAL DISTRIBUTION

$$\text{Mean} = E(X) = \mu'_1 = \int_0^\infty x f(x)\,dx = \lambda \int_0^\infty x e^{-\lambda x}\,dx$$

$$= \lambda \int_0^\infty e^{-\lambda x} x^{2-1}\,dx = \lambda \frac{\Gamma 2}{\lambda^2} = \frac{1}{\lambda} \underline{|1} \qquad \left\{ \text{Since } \int_0^\infty e^{-\alpha x} x^{n-1}\,dx = \frac{\Gamma n}{\alpha^n} \right\}$$

$$\Rightarrow \qquad \boxed{\text{Mean} = E(X) = \mu'_1 = \frac{1}{\lambda}}$$

Now, $$\mu'_2 = \int_0^\infty x^2 f(x)dx = \lambda \int_0^\infty x^2 e^{-\lambda x} dx$$

$$= \lambda \int_0^\infty e^{-\lambda x} x^{3-1} dx = \lambda . \frac{\Gamma 3}{\lambda^3}$$

$$= \frac{1}{\lambda^2} \lfloor 2 \quad \left\{ \text{Since } \Gamma n+1 = \lfloor n \right\}$$

$$\Rightarrow \quad \mu'_2 = \frac{2}{\lambda^2}$$

Hence, $$\text{variance} = \sigma^2 = \mu'_2 - {\mu'_1}^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$\Rightarrow \quad \boxed{\text{variance} = \sigma^2 = \frac{1}{\lambda^2}}$$

Hence, Standard Deviation $= \sqrt{\text{Variance}} = \frac{1}{\lambda}$

MOMENT GENERATING FUNCTION OF EXPONENTIAL DISTRIBUTION

$$\boxed{M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-1}}$$

MEAN AND VARIANCE OF EXPONENTIAL DISTRIBUTION WITH THE HELP OF MGF

MGF of exponential distribution is given as

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-1} = 1 + \frac{t}{\lambda} + \frac{t^2}{\lambda^2} + \ldots..$$

$$\Rightarrow \quad \text{coefficient of } \frac{t}{\lfloor 1} = \mu_1' = \frac{1}{\lambda}$$

and $\quad \text{coefficient of } \dfrac{t^2}{\lfloor 2} = \mu_2' = \dfrac{2}{\lambda^2}$

$$\Rightarrow \quad \boxed{\text{mean} = \mu_1' = \frac{1}{\lambda}}$$

and $\quad \text{variance} = \mu_2' - \mu_1'^2$

$$\Rightarrow \quad \boxed{\text{variance} = \frac{1}{\lambda^2}}$$

MEMORY LESS PROPERTY OF EXPONENTIAL DISTRIBUTION

The most important property of exponential distribution is its "Memory less" property, which means that if lifetime of an item is exponentially distributed then an item which has been in use for some hours is as good as a new item with regard to the amount of time left untill the item lasts.

A exponentially distributed random variable X is called memory less if,

$$P\left(\frac{X > x_1 + t}{X > t}\right) = P(X > x_1), \ \forall\, x_1, t > 0$$

Now, $$P\left(\frac{X > x_1 + t}{X > t}\right) = \frac{P(X > x_1 + t \cap X > t)}{P(X > t)} = \frac{P(X > x_1 + t)}{P(X > t)}$$

$$= \frac{1 - P(X \leq x_1 + t)}{1 - P(X \leq t)} = \frac{1 - [1 - e^{-\lambda(x_1 + t)}]}{1 - [1 - e^{-\lambda t}]}$$

$$= e^{-\lambda x_1} = 1 - (1 - e^{-\lambda x_1}) = 1 - P(X \leq x_1)$$

$$= P(X > x_1)$$

Hence, $$\boxed{P\left(\frac{X > x_1 + t}{X > t}\right) = P(X > x_1)}$$

Q.1. *Show that for the exponential distribution given by $dp = ae^{-x/c}\,dx$; $0 \le x < \infty, c > 0$, a being a constant, the mean and standard deviation are each equal to c.*

Sol. Given $dp = ae^{-x/c}\,dx\,;\, 0 \le x < \infty$

Since $\int_0^\infty dp = 1 \Rightarrow \int_0^\infty ae^{-\frac{x}{c}}\,dx = 1$

$\Rightarrow \quad -ac\left(e^{-x/c}\right)_0^\infty = 1 \Rightarrow -ac\,(0-1) = 1 \Rightarrow a = \frac{1}{c}$

We have $\mu'_r = E(X^r) = \int_0^\infty x^r a e^{-\frac{x}{c}}\, dx = \int_0^\infty c^r u^r a e^{-u}\, c\, du$

$$\left[\text{Put } \frac{x}{c} = u \Rightarrow x = cu \Rightarrow dx = c\,du\right]$$

$$= c^r \int_0^\infty e^{-u}\, u^{(r+1)-1}\, du = c^r \,\overline{\left| r+1 \right.}$$

$\Rightarrow$ $\mu'_r = c^r \,\underline{\left| r \right.}$

$\Rightarrow$ $\text{Mean} = \mu'_1 = c$

and $\text{Variance} = \mu_2 = \mu'_2 - {\mu'_1}^2 = c^2 \,\overline{\left| 3 \right.} - c^2 = 2c^2 - c^2 = c^2$

$\Rightarrow$ $\text{standard deviation} = \sqrt{c^2} = c$

Q. 2. *The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = \frac{1}{2}$.*

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

(b) *What is the conditional probability that a repair takes atleast 10 hours given that its duration exceeds 9 hours?*

Sol. Let random variable X represents time (in hours) required to repair a machine, then its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{2} e^{-\frac{x}{2}} & , \quad x \geq 0 \\ 0 & , \quad \text{otherwise} \end{cases}$$

(a)
$$P(X > 2 \text{ hours}) = \int_2^\infty f(x)dx = \frac{1}{2} \int_2^\infty e^{-x/2} dx$$

$$= \frac{1}{2} \left(-2 e^{-x/2}\right)_2^\infty$$

$$= -(0 - e^{-1})$$

$$= e^{-1}$$

$$= 0.3679$$

(b) $$P\left(\frac{X \geq 10}{X > 9}\right) = P(X > 1) \text{ (using memory less property)}$$

$$= \int_1^\infty f(x)\,dx \qquad = \frac{1}{2}\int_1^\infty e^{-\frac{x}{2}}\,dx$$

$$= \frac{1}{2}\left(-2e^{-\frac{x}{2}}\right)_1^\infty \qquad = -\left(e^{-x/2}\right)_1^\infty$$

$$= -(0 - e^{-1/2}) \qquad = e^{-1/2}$$

$$= 0.6065$$

Exercise

1. The probability of a man hitting a target is $\frac{1}{4}$. How many times must he fire so that the probability of hitting the target at least once is greater than $\frac{2}{3}$?

2. Five fair coins were tossed 100 times. Calculate expected frequencies using the following outcomes:

No. of heads :	0	1	2	3	4	5
Frequencies :	2	10	24	35	18	8

3. A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantees that not more than10 pins will be defective, then what is the approximate probability that a box will fail to meet the guaranteed quality?

4. A skilled typist, on routine work, kept a record of mistakes made per day during 300 working days as follows:

Mistakes per day :	0	1	2	3	4	5	6
No. of days :	143	90	42	12	9	3	1

Fit a Poisson distribution to the above data and hence calculate the theoretical frequencies.

5. In a normal distribution, 7% of the items are under 35 and 89% are under 63. What is the mean and standard deviation of the distribution?

6. Fit a normal curve to the following frequency distribution:

x	:	4	6	8	10	12	14	16	18	20	22	24
f	:	1	7	15	22	35	43	38	20	13	5	1

7. Subway trains from Karolbagh to Chandani Chowk run every half an hour between midnight and six in the morning, What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes?

8. The mileage which car owner gets with a certain kind of radial tyre is a random variable having an exponential distribution with mean 40000 km. Find the probability that one of these types will last at least 20000 km.

Answers

1. 4
2. 3.125, 15.625, 31.25, 15.625, 3.125
3. $1-\sum_{x=0}^{10}\frac{(0.0067385)5^x}{\underline{|x}}$
4. 123, 110, 49, 14, 3, 1, 0
5. 50.29, 10.33
6. $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, where $\mu=13.85, \sigma=3.83$
7. $\frac{1}{3}$
8. 0.6065

THANKS