PROBABILITY AND STATISTICS (UCS401)

Lecture-22

(Normal distribution with illustrations)

Random Variables and their Special Distributions(Unit –III & IV)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Dr. Rajanish Rai
Assistant Professor
School of Mathematics
Thapar Institute of Engineering and Technology, Patiala

Normal distribution

A continuous random variable X which has the following probability density function (p.d.f.)

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

is called normal variate and its distribution is called normal distribution.

Or

A random variable X is said to follow normal distribution with parameter μ (called mean) and σ^2 (called variance) if its probability density function is given by

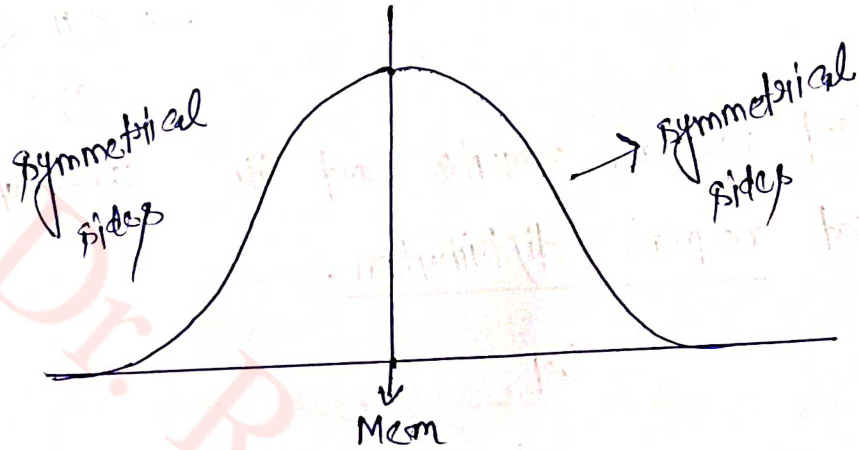
$$f(x;\mu,\sigma^2) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & -\infty < x < \infty \\ & -\infty < \mu < \infty \\ 0 & 0 < \sigma < \infty \end{cases}$$

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \, ;$$

$E(X) = \mu$ & $Var(X) = \sigma^2$

For Convinience, we denote it as $X \sim N(\mu, \sigma^2)$ read as X follows normal distribution with mean μ and variance σ^2.

Normal distribution is also known as Gaussian distribution or Bell-shaped curve distribution.



Properties of normal distribution :-

(i) The mean, median and mode are equal.

(ii) The curve is symmetric about the mean

(iii) Total area under the curve is 1.

(iv) Area to the left and area to the right about the mean are same, i.e, 0.5.

Standard normal variate :-

For normal distribution $X \sim N(\mu, \sigma^2)$ the variable $Z = \frac{X-\mu}{\sigma}$ is called standard normal variate with mean=0 and standard deviation = 1.

$\because$ The p.d.f. of X is

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

The p.d.f. of Z is

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

Mean of $Z = E(Z) = E\left(\frac{X-\mu}{\sigma}\right)$

$\because E(aX) = a\,E(X)$

$E(X+Y) = E(X) + E(Y)$

$$= \frac{1}{\sigma} E(X-\mu)$$

$$= \frac{1}{\sigma}\left[E(X) - E(\mu)\right]$$

$$= \frac{1}{\sigma}\left[\mu - \mu\right]$$

$$\boxed{E(Z) = 0}$$

$\because V(\alpha X) = \alpha^2 Var(X)$

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right)$$

$$= \frac{1}{\sigma^2} Var(X-\mu)$$

$$= \frac{1}{\sigma^2} Var(X)$$

$$= \frac{1}{\sigma^2}\sigma^2 = 1$$

$$\boxed{Var(Z) = 1}$$

$$Var(X) = E\left(X - E(X)\right)^2$$

$$Var(X+a) = E\left(X+a - E(X+a)\right)^2$$

$$= E\left(X+a - E(X) - E(a)\right)^2$$

$$= E\left(X+a - E(X) - a\right)^2$$

$$= E\left(X - E(X)\right)^2 = Var(X)$$

$$Var(X+a) = Var(X)$$

Mean and variance of normal distribution :-

Mean :- The probability density function is given by

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

$$\text{Mean} = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

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

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

put $\frac{x-\mu}{\sigma} = z \Rightarrow x = \mu + \sigma z$

$\Rightarrow dx = \sigma dz$

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

$$= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz + \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} \cancel{z e^{-z^2/2} dz}}_{0}$$

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

put $\frac{z^2}{2} = u$

$z dz = du$

$dz = \frac{du}{\sqrt{2u}}$

$$= \frac{2\mu}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-u} \frac{du}{\sqrt{2u}}$$

$$= \frac{2\mu}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} \int_0^\infty e^{-y}\, y^{-1/2}\, dy$$

$$\Gamma n = \int_0^\infty e^{-x} x^{n-1} dx$$

$$\Gamma \tfrac{1}{2} = \sqrt{\pi}$$

$$\Gamma n+1 = n \Gamma n$$

$$= \frac{\mu}{\sqrt{\pi}} \int_0^\infty e^{-y}\, y^{1/2-1}\, dy$$

$$= \frac{\mu}{\sqrt{\pi}} \Gamma \tfrac{1}{2} = \frac{\mu}{\cancel{\sqrt{\pi}}} \cancel{\sqrt{\pi}} = \mu$$

$$\boxed{\text{Mean} = E(X) = \mu}$$

(ii) Variance :- The variance is defined as :

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

$$\therefore \quad E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

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

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

put $\frac{x-\mu}{\sigma} = z \Rightarrow x = \mu + \sigma z \Rightarrow dx = \sigma dz$

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

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

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

$$\left. + \; 2\mu\sigma \cancelto{0}{\int_{-\infty}^{\infty} z e^{-z^2/2}\, dz} \right]$$

$$= \frac{1}{\sqrt{2\pi}} \left[2\mu^2 \int_{0}^{\infty} e^{-z^2/2}\, dz + 2\sigma^2 \int_{0}^{\infty} z^2 e^{-z^2/2}\, dz \right]$$

let $\frac{z^2}{2} = p \Rightarrow dz = \frac{dp}{\sqrt{2p}}$

$$= \frac{2}{\sqrt{2\pi}} \left[\mu^2 \int_0^{\infty} e^{-p} \frac{dp}{\sqrt{2p}} + \sigma^2 \int_0^{\infty} (2p) e^{-p} \frac{dp}{\sqrt{2p}} \right]$$

$$= \frac{\sqrt{2}}{\sqrt{\pi}} \left[\frac{\mu^2}{\sqrt{2}} \left(\int_0^{\infty} e^{-p} p^{1/2 - 1}\, dp \right) + \frac{2\sigma^2}{\sqrt{2}} \left(\int_0^{\infty} e^{-p} p^{3/2 - 1}\, dp \right) \right]$$

$$= \frac{\sqrt{2}}{\sqrt{\pi}} \left[\frac{\mu^2}{\sqrt{2}} \Gamma\tfrac{1}{2} + \frac{2\sigma^2}{\sqrt{2}} \Gamma\tfrac{3}{2} \right] \qquad \Gamma n = \int_0^{\infty} e^{-x} x^{n-1}\, dx$$

$$= \frac{\sqrt{2}}{\sqrt{\pi}} \left[\frac{\mu^2}{\sqrt{2}} \sqrt{\pi} + \frac{2\sigma^2}{\sqrt{2}} \frac{1}{2} \sqrt{\pi} \right] \qquad \Gamma\tfrac{1}{2} = \sqrt{\pi}$$

$$\Gamma(n+1) = n\,\Gamma n$$

$$= \frac{\sqrt{2}}{\sqrt{\pi}} \frac{\sqrt{\pi}}{\sqrt{2}} \left[\mu^2 + \sigma^2 \right]$$

$$\boxed{E(X^2) = \mu^2 + \sigma^2}$$

Now, $Var(X) = E(X^2) - (E(X))^2 = (\mu^2 + \sigma^2) - \mu^2$

$$\boxed{Var(X) = \sigma^2}$$

(3) Moment generating function:-

The probability density function for normal distribution is given by—

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

The moment generating function is given by

$$M_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

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

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

put $\frac{x-\mu}{\sigma} = z \Rightarrow x = \mu + \sigma z \; ; \; dx = \sigma dz$

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

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

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(z^2 - 2\sigma z t + \sigma^2 t^2 - \sigma^2 t^2\right)} dz$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-\sigma t)^2} e^{\frac{\sigma^2 t^2}{2}} dz .$$

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

$$= \frac{2e^{\mu t+\frac{\sigma^2 t^2}{2}}}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{1}{2}(z-\sigma t)^2} dz$$

put $\frac{1}{2}(z-\sigma t)^2 = \theta$

$$(z-\sigma t)\, dz = d\theta$$

$$dz = \frac{d\theta}{(z-\sigma t)} = \frac{d\theta}{\sqrt{2\theta}}$$

$$= \frac{2e^{\mu t+\frac{\sigma^2 t^2}{2}}}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\theta} \frac{d\theta}{\sqrt{2\theta}}$$

$$= \frac{2e^{\mu t+\frac{\sigma^2 t^2}{2}}}{\sqrt{2}\sqrt{2\pi}} \int_{0}^{\infty} e^{-\theta} \theta^{\frac{1}{2}-1} d\theta$$

$$= \frac{1}{\sqrt{\pi}} e^{\mu t+\frac{\sigma^2 t^2}{2}} \Gamma\tfrac{1}{2} = \frac{1}{\cancel{\sqrt{\pi}}} e^{\mu t+\frac{\sigma^2 t^2}{2}} \cancel{\sqrt{\pi}}$$

$$\boxed{M_X(t) = e^{\mu t+\sigma^2 t^2/2}}$$

Moment generating function of standard normal variate :

$\because$ For $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \Rightarrow M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$.

If $Z = \frac{X-\mu}{\sigma} \Rightarrow f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \Rightarrow M_Z(t) = ?$

$$M_Z(t) = E\left(e^{tZ}\right) = E\left(e^{\left(\frac{X-\mu}{\sigma}\right)t}\right)$$

$$= E\left(e^{\frac{Xt}{\sigma}} e^{-\mu t/\sigma}\right)$$

$$= e^{-\mu t/\sigma} E\left(e^{(t/\sigma)X}\right)$$

$$= e^{-\mu t/\sigma} M_X(t/\sigma)$$

$\because M_X(t) = e^{\mu t + \sigma^2 t^2/2} \Rightarrow M_X(t/\sigma) = e^{\frac{\mu t}{\sigma} + \frac{\sigma^2}{2}\left(\frac{t}{\sigma}\right)^2} = e^{\frac{\mu t}{\sigma} + t^2/2}$

$$= e^{-\mu t/\sigma} e^{(\mu t/\sigma) + t^2/2}$$

$\boxed{M_Z(t) = e^{t^2/2}}$; $Z = \frac{X-\mu}{\sigma}$.

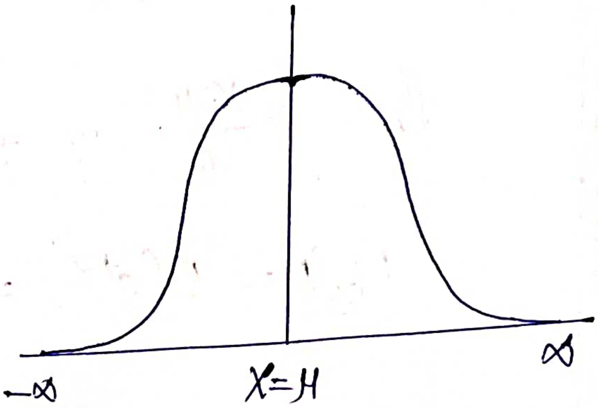
$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \Rightarrow M_Z(t) = e^{t^2/2}$ Mean $= 0$, Variance $(\sigma^2) = 1$

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \Rightarrow M_X(t) = e^{\mu t + \sigma^2 t^2/2}$ Mean $= \mu$, Variance $= \sigma^2$

Normal distribution — Area under the Curve

The probability density function for normal distribution is given by -

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



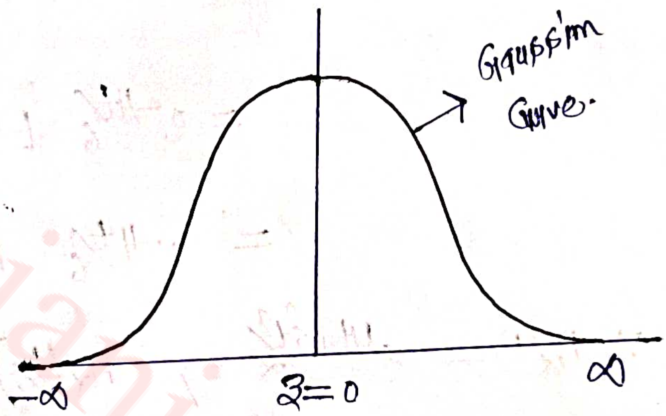
For Normal variate

$$Z = \frac{X-\mu}{\sigma}$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \; ; \; z \in \mathbb{R}$$

Mean $= 0$

Variance $(\sigma^2) = 1$



Question :- If the height of 300 students are normally distributed with mean 64.5 inch and standard deviation 3.3 inch. How many students have height

(i) less than 5 feets.

(ii) between 5 feets and 5 feet 9 inches.

solution :- Given that

$$\mu = 64.5 \quad \text{and} \quad N = 300$$

$$\sigma = 3.3$$

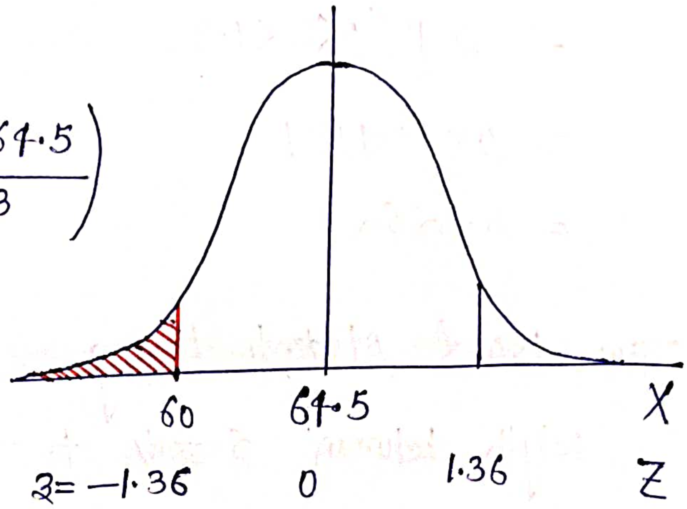
$$Z = \frac{X-\mu}{\sigma} = \frac{X-64.5}{3.3}$$

(i) less than 5 feets = 60 inches

Thus, required probability

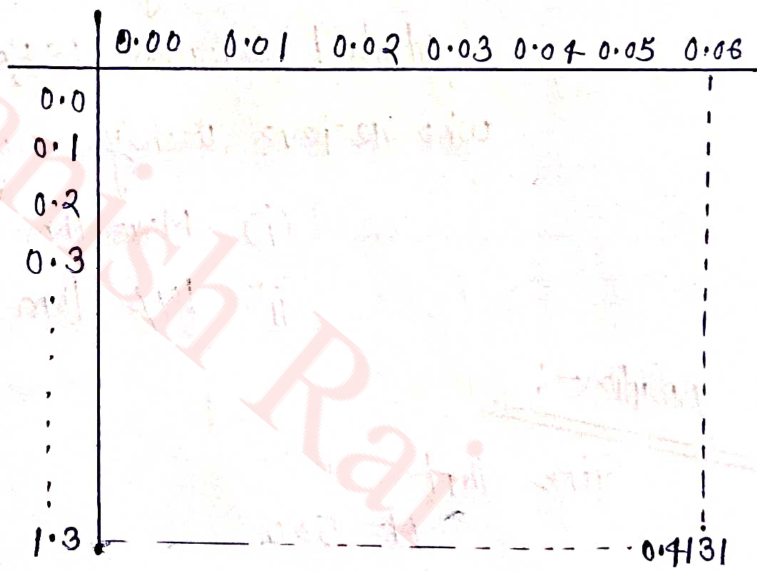
$$P(X < 60) = P\left(Z < \frac{60-64.5}{3.3}\right)$$

$$= P(Z < -1.36)$$



$$= 0.5 - P(-1.36 < Z < 0)$$

$$= 0.5 - P(0 < Z < 1.36) \rightarrow \text{by symmetry}$$



$$= 0.5 - 0.4131$$

$$= 0.0869.$$

Thus No. of students having heigth less than 5 feets $= 300 \times 0.0869$

$= 26.07 \approx 26$ students Ans

(ii) between 5 feets to 5 feet 9 inches $\quad Z = \frac{X - 64.5}{3.3}$

$$P(60 < X < 69) = P\left(\frac{60-64.5}{3.3} < Z < \frac{69-64.5}{3.3}\right)$$

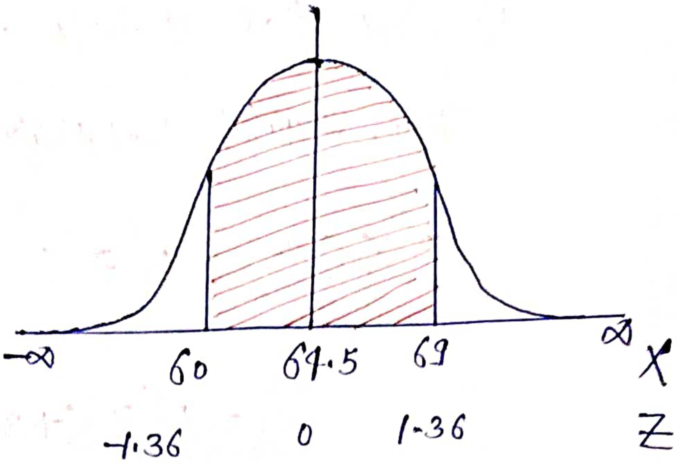
$$= P(-1.36 < Z < 1.36)$$

$= P(-1.36 < Z < 1.36)$

$= 2\,P(0 < Z < 1.36)$

$= 2 \times 0.4131$

$= 0.8262.$



Thus, No. of students having heigth between 5 feets to 5 feets 9 inches

$= 300 \times 0.8262 = 247.8 \approx 248$ **Ans**

Question:- The distribution of 500 workers in a factory is approximately normal with mean = 75 Rs and standard deviation 15 Rs. Find the number of workers who recieve weekly wages:

(i) More than 90

(ii) Less than 45.

	0.00	0.01
0.0		
0.1		
0.2		
⋮		
1.0	0.3413	
⋮		
2.0	0.4772	

Solution:-

Given that

$N = 500$

$\mu = 75$ & $\sigma = 15$

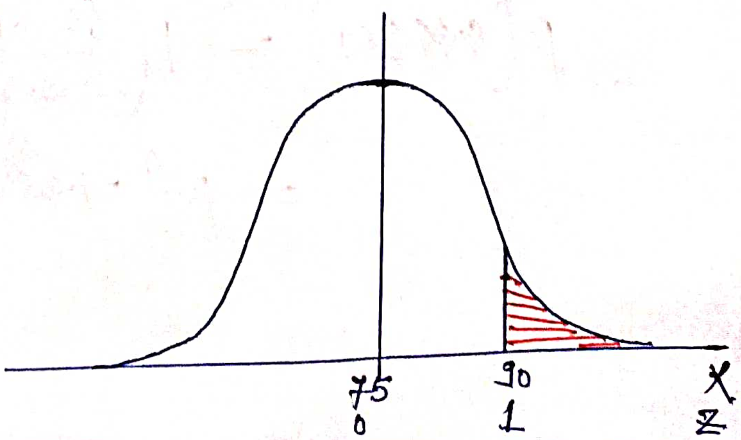
$Z = \dfrac{x-\mu}{\sigma} = \dfrac{x-75}{15}$

Thus, required probability

$= P(X > 90)$

$= P\left(Z > \dfrac{90-75}{15}\right)$

$= P(Z > 1)$



$= P(Z > 1)$

$= 0.5 - P(0 < Z < 1)$

$= 0.5 - 0.3413$

$$\boxed{P(X > 90) = 0.1587}$$

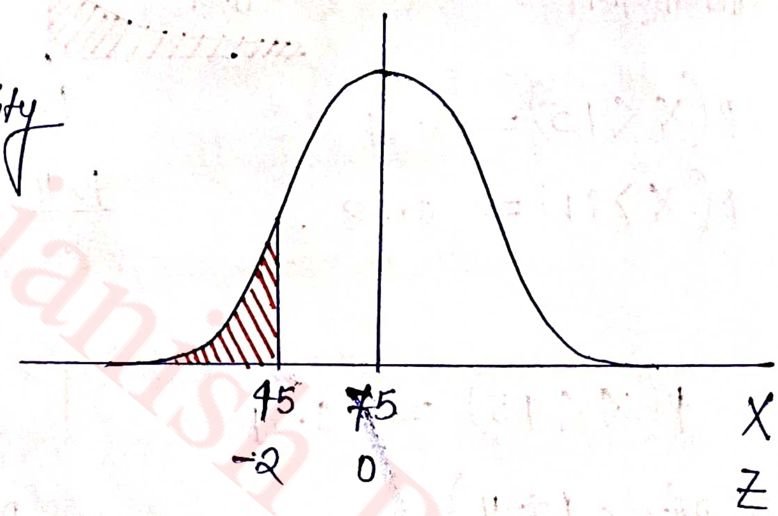
Thus, No. of workers who recieve weekly wages more than 90 Rs

$= 500 \times 0.1587$

$= 79.35 \approx 79$ Ans

(ii) Less than 45

Thus required probability



$P(X < 45)$

$= P\left(Z < \frac{45-75}{15}\right)$

$= P(Z < -2)$

$= 0.5 - P(-2 < Z < 0)$

$= 0.5 - P(0 < Z < 2)$ → by symmetry

$= 0.5 - 0.4772$

$= 0.0228.$

∴ No. of workers whose weekly wages is less than 45

$= 500 \times 0.0228$

$= 11.4 \approx 11$ workers Ans

Question:- In normal distribution, 31% of the items are under 45 and 8% are over 64. Find mean and standard deviation of the distribution. Given that

$$p(Z<-1.4)=0.08$$

and $$p(Z>0.5)=0.31.$$

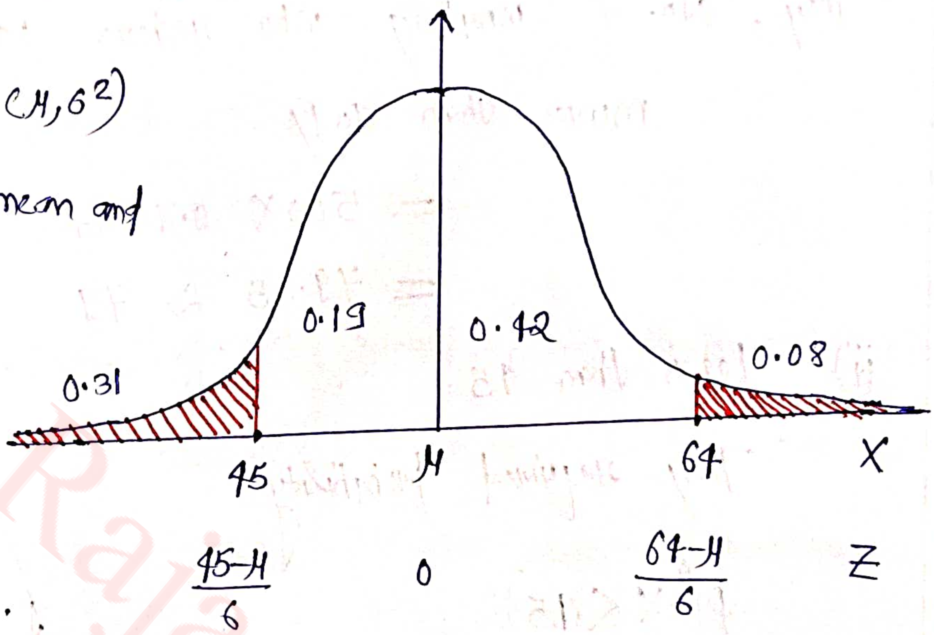
Solution:- Let $X \sim N(\mu, \sigma^2)$

Let μ and σ be mean and S.D. of X

Given that

$p(X<45)=0.31$

$p(X>64)=0.08$.



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

$$p\left(Z<\frac{45-\mu}{\sigma}\right)=0.31$$

Given that

$$p(Z>0.5)=0.31$$

$$\Rightarrow p(Z<-0.5)=0.31$$

$$\Downarrow$$

$$\frac{45-\mu}{\sigma}=-0.5 \quad \text{—(I)}$$

$$\because p(X>64)=0.08$$

$$p\left(Z>\frac{64-\mu}{\sigma}\right)=0.08$$

Given that

$$p(Z<-1.4)=0.08$$

$$p(Z>1.4)=0.08$$

$$\Downarrow$$

$$\frac{64-\mu}{\sigma}=1.4 \quad \text{—(II)}$$

By deviding, we get

$$\frac{75-\mu}{64-\mu} = -\frac{0.5}{1.7}$$

$\Rightarrow \quad \mu = 50$

$(2) \Rightarrow \quad \sigma = 10$

Ans

Question:- If skill are classified as A, B & C according as the length, breath, index as under 75, between 75 and 80 or over 80, find the approximately the mean and standard devition of a series in which A are 58%, B are 38% and C are 4%, being given that

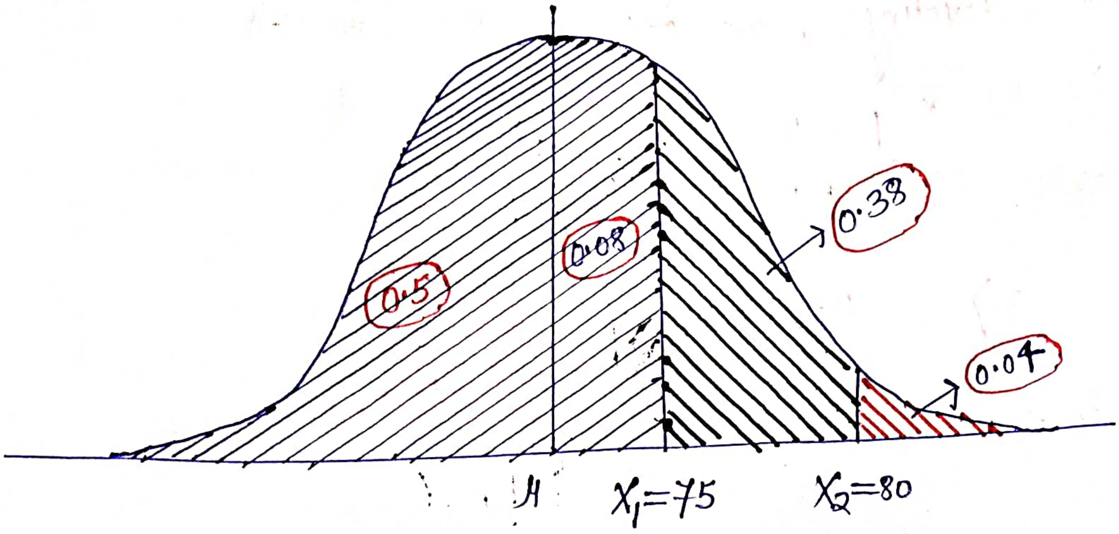
$$P(0<Z<0.2) = 0.08$$
$$P(0<Z<1.75) = 0.46.$$

Solution:- Let $X \sim N(\mu, \sigma^2)$

Let μ and σ be mean and S.D. of X.

Given that

A	$P(X<75)$	0.58
B	$P(75<X<80)$	0.38
C	$P(X>80)$	0.04



$$Z_1 = \frac{X_1-\mu}{\sigma} \qquad Z_2 = \frac{X_2-\mu}{\sigma}$$

$$Z_1 = \frac{75-\mu}{\sigma} \qquad Z_2 = \frac{80-\mu}{\sigma}$$

Given $P(0<Z<Z_1) = 0.08$

$P(0<Z<Z_2) = 0.08 + 0.38 = 0.46$

$P(0<Z<Z_1) = 0.08$

Given that

$P(0<Z<0.20) = 0.08$

$\Downarrow$

$Z_1 = 0.2$

$P(0<Z<Z_2) = 0.46$

given that $P(0<Z<1.75) = 0.46$

$\Downarrow$

$Z_2 = 1.75$

$$0.2 = \frac{75-\mu}{\sigma}$$

$\Downarrow$

$\mu + (0.2)\sigma = 75$ — (1)

$$1.75 = \frac{80-\mu}{\sigma}$$

$\mu + (1.75)\sigma = 80$ — (2)

solving (1) and (2), we get

$\boxed{\mu = 74.4}$ and $\boxed{\sigma = 3.5}$ Ans

Question :- 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 one of them would have scored above 75? Given that

$$P(Z<-2) = 0.0228.$$

Solution :- X : No. of students whose marks are above 75.

$n = 3$

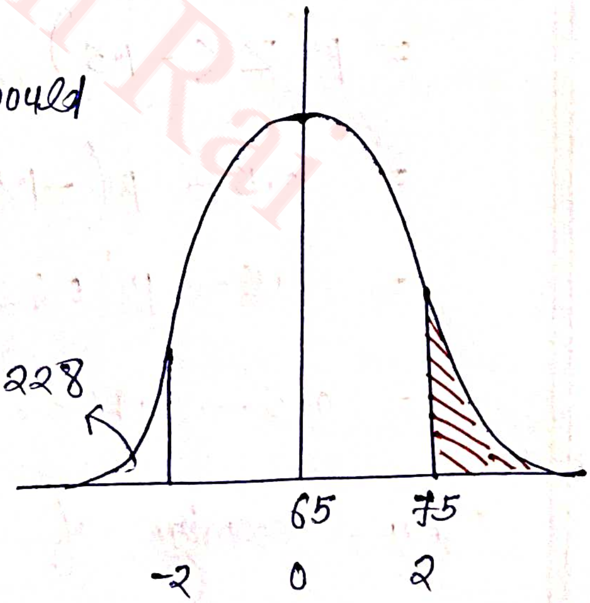
$p = P(X'>75)$

X' : be the marks obtained by a ~~student~~ students in a subject,

Given $\mu = 65$, $\sigma = 5$

$$Z = \frac{X'-\mu}{\sigma} = \frac{X'-65}{5}$$

The probability that a students would have scored above 75 is



$= P(X'>75)$

$= P\left(Z > \frac{75-65}{5}\right)$

$= P(Z>2)$

$= \cancel{0.5 - P(0<Z<2)}$

$= P(Z<-2) \rightarrow$ by symmetry

$= 0.0228.$

Thus $n = 3$ (finite) $\longrightarrow$ Binomial distribution.

$$p = P(x' > 75) = 0.0228$$

$$q = 1 - p = 1 - 0.0228 = 0.9772$$

Since 3 students are selected at random from given group, thus required probability that at least one of them would have scored above 75 is

$$P(X \geq 1)$$

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

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

$$= 1 - {}^3C_0\, p^0 q^3$$

$$= 1 - q^3$$

$$= 1 - (0.9772)^3$$

$$= 1 - 0.9331$$

$$= 0.066685$$

$$= 0.0669$$ Ans

$$P(X \geq 1) = 0.0669$$ Ans