Date ___/___/___

$$= \frac{1}{b-a}\left[\frac{x^{r+1}}{r+1}\right]_a^b$$

$$= \frac{1}{b-a}\left[\frac{b^{r+1}-a^{r+1}}{r+1}\right]$$

$$\boxed{\text{Mean} = \mu_1' = \frac{b^2-a^2}{(b-a)2} = \frac{a+b}{2}}$$

$$\mu_2' = \frac{b^3-a^3}{(b-a)3} = \frac{a^2+ab+b^2}{3}$$

$$\text{variance} = \mu_2 = \mu_2' - (\mu_1')^2$$

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

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

us A Rectangular Distribution is defined as

$$f(x) = \begin{cases} \frac{1}{2a} & \text{if } -a \le x \le a \\ 0 & \text{otherwise} \end{cases}$$

find the moment generating function and μ_{2r}, μ_{2r+1} ?

Date __/__/____

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

$$= \int_{-a}^{a} e^{tx} f(x)\,dx$$

$$= \int_{-a}^{a} e^{tx} \frac{1}{2a}\,dx = \frac{1}{2a}\left\{\frac{e^{tx}}{t}\right\}_{-a}^{a}$$

$$= \frac{1}{2at}\left\{e^{at} - e^{-at}\right\}$$

$$= \frac{1}{at} \sinh at$$

$$M_x(t) = \frac{1}{at}(\sinh at)$$

$$= \frac{1}{at}\left\{(at) + \frac{(at)^3}{\lfloor 3} + \frac{(at)^5}{\lfloor 5} + \cdots\cdots\right\}$$

$$= \left\{1 + \frac{a^2t^2}{\lfloor 3} + \frac{t^4a^4}{\lfloor 5} + \cdots\cdots\right\}$$

$\mu_r' =$ Cofe. of $\frac{t^r}{\lfloor r}$ in $M_x(t)$

$\mu_1' = \mu_3' = \mu_5' = \cdots\cdots = 0$

$\mu_1' = 0 \quad = $ Mean

$\mu_r = \mu_r'$

$\mu_{2r+1} = 0$

$$\text{Coff. of } \frac{t^{2r}}{\lfloor 2r} = \mu'_{2r} = \frac{a^{2r}\lfloor 2r}{\lfloor 2r+1}$$

$$= \frac{a^{2r}}{2r+1} = \mu_{2r}$$

Ques

$$M_{(x-\mu)}(t) = E(e^{t(x-\mu)})$$

$$= \int_{-\infty}^{\infty} e^{t(x-\mu)} f(x)\,dx$$

$$= e^{-t\mu}\int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$\boxed{M_{(x-\mu)}(t) = e^{-t\mu} M_x(t)}$$

Ques The mean and variance of a continuous uniform distribution are 1 and $4/3$ Respectively find $P(X<1.5)$?

Ans

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Mean} = \frac{a+b}{2} = 1 \quad \text{--- (i)}$$

$$\text{variance} = \frac{(b-a)^2}{12} = 4/3 \quad \text{--- (ii)}$$

Date ___ / ___ / ___

$a + b = 2$

$b - a = 4$

$a + a + 4 = 2$

$2a = -2$

$\boxed{a = -1}$

$\boxed{b = 3}$

$$f(x) = \begin{cases} \frac{1}{4} & -1 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

$$P(X < 1.5) = \int_{-\infty}^{1.5} f(x)\,dx$$

$$= \int_{-1}^{1.5} \frac{1}{4}\,dx$$

$$= \frac{1}{4}(x)_{-1}^{1.5}$$

$$= \frac{2.5}{4}$$

Exponential distribution :-

A random variable x is said to have an exponential distribution with parameter $\alpha > 0$, if its P.d.f is given by:

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

The Cumulative distribution function $F(x)$ is given by

$$F(x) = \int_0^x f(t)\,dt = \int_0^x \alpha e^{-\alpha t}\,dt \quad \{\text{when } x \geq 0$$

$$= \alpha \left\{ \frac{e^{-\alpha t}}{-\alpha} \right\}_0^x$$

$$= (1 - e^{-\alpha x} \quad \text{if } x \geq 0$$

$$\therefore \quad F(x) = \begin{cases} 1 - e^{-\alpha x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Moment generating function or Exponential distribution :-

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

$$= \int_0^\infty e^{tx} \alpha e^{-\alpha x}\,dx$$

$$= \alpha \int_0^\infty e^{-(\alpha - t)x}\,dx \quad \text{where } \alpha > t$$

Date __/__/____

$$= \alpha \left[\frac{e^{-(\alpha - t)x}}{-(\alpha - t)} \right]_0^\infty$$

$$= \alpha \left[0 + \frac{1}{\alpha - t} \right]$$

$$= \frac{\alpha}{\alpha - t}$$

$$= \left(1 - \frac{t}{\alpha}\right)^{-1}$$

$$= 1 + \frac{t}{\alpha} + \frac{t^2}{\alpha^2} + \cdots \quad \text{where } \alpha > t$$

$$= \sum_{r=0}^{\infty} (t/\alpha)^r$$

$\mu_r' =$ rth moment about origin

$= \text{Coff. of } \frac{t^r}{\lfloor r} \text{ in } M_X(t)$

$$= \frac{r!}{\alpha^r}$$

$$\mu_1' = \frac{1}{\alpha}$$

$$\mu_2' = \frac{3}{\alpha^2}$$

$$\text{mean} = \mu_1' = \frac{1}{\alpha}$$

$$\text{var}(x) = \mu_2 = \mu_2' - \mu_1'^2 = \frac{2}{\alpha^2} - \frac{1}{\alpha^2} = \frac{1}{\alpha^2}$$

Saathi

Date ___/___/___

Hence if $x \sim exp(\alpha)$, then mean $= 1/\alpha$ and variance $= \frac{1}{\alpha^2}$ and Standard deviation $= \frac{1}{\alpha}$

Memoryless property of Exponential Distribution

If x is Exponentially distributed then

$P(X > s+t \mid x > s) = P(X > t)$ for any $s, t > 0$

we know that

$$P(X > K) = \int_K^\infty f(x)dx = \int_K^\infty \lambda e^{-\lambda x} dx = e^{-\lambda K}$$

~~$\therefore P(x > K) =$~~

$$\therefore P(X > s+t \mid x > s) = \frac{P(X > s+t \text{ and } X > s)}{P(X > s)}$$

$$= \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

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

$$= P(X > t)$$

Date __ / __ / ____

The converse of above result is also true.
Hence if, $P(X > s+t \mid X > s)$ then X follows Exponential distribution.

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

(i) What is the probability that the repair time Exceeds 2 hrs?

(ii) What is the conditional Probability that a repair takes at least 10 hours given that its duration exceeds 9 hours?

Soluⁿ Here X represents time (in hours) required to repair a machine, than its pdf is given as

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

(i) $$P(X > 2) = \int_0^{\infty} f(x)\,dx$$

$$= \frac{1}{2} \int_2^{\infty} e^{-x/2}\,dx = \frac{1}{2} \left[\frac{e^{-x/2}}{-1/2} \right]_2^{\infty}$$

$$= e^{-1} = 0.3679$$

(ii) $P(x \geq 10 / x > 9) = P(x > 1)$ (By memoryless property)

Date ___/___/___

$$= \int_1^\infty f(x)\,dx$$

$$= \frac{1}{2}\int_1^\infty e^{-x/2}\,dx$$

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

Ques- If X is Exponentially distributed with parameter λ, prove that the probability that X exceeds its Expected value is less than 0.5.

Solⁿ Let X is Exponentially distributed with parameter λ hence its pdf is

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

Also we Expected value of $X = E(X) = \text{mean} = 1/\lambda$

Required Probability $= P(X \geq 1/\lambda)$

$$= \int_{1/\lambda}^\infty f(x)\,dx$$

$$= \lambda \int_{1/\lambda}^\infty e^{-\lambda x}\,dx$$

$$= -\left(e^{-\lambda x}\right)_{1/\lambda}^\infty$$

Date __/__/__

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

$= e^{-1}$

$= 0.366 < 0.5$

Hence $P(x \geq 1/4) < 0.5$

Ques The mileage which car owner gets with a certain kind of radial tyre is a random variable having an exponential distribution with mean 40,000 km. find the probability that one of these types will last (i) at least 20,000 km
(ii) at most 30,000 km

Soln Let random variable x denote the mileage obtaine with pdf $f(x) = \frac{1}{40,000} e^{-x/40,000}, x \geq 0$

$\{ \because mean = 1/\lambda$

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

$= e^{-0.5} = 0.6065$

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

$= (e^{-3/4} - 1) = 0.5276$

Date __ / __ / __

Normal distribution :- A random variable 'x' is said to follow normal distribution if its probability mass function is given by

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

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

$$N(\mu, \sigma^2)$$

Ques 1. Show that function $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$
$\sigma > 0$, $-\infty < x < \infty$, $-\infty < \mu < \infty$

Ans.
(i) $\sigma > 0$ (given)

we know that $e^{-f\left(\frac{x-\mu}{\sigma}\right)^2} > 0, \forall x$

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

$$f(x) > 0 \quad \forall x$$

(ii) $\int_{-\infty}^{\infty} f(x)\,dx$

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

Date __/__/___

let $\frac{x-\mu}{\sigma} = t$

$x = \sigma t + \mu$

$dx = \sigma\, dt$

$$I = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2} \sigma\, dt$$

$$I = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2}\, dt$$

$t^2/2 = u \Rightarrow t = \sqrt{2}\sqrt{u}$

$dt = \frac{\sqrt{2}}{2\sqrt{u}}\, du \quad dt = \frac{\sqrt{2}}{2\sqrt{u}}\, du$

$$I = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-u} \frac{\sqrt{2}}{2\sqrt{u}}\, du$$

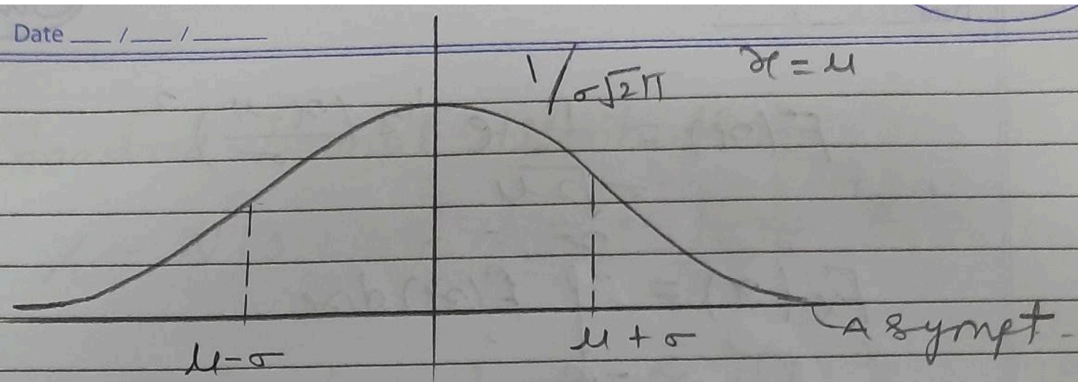
$$I = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-u} \frac{1}{\sqrt{u}}\, du$$

$$= \frac{1}{\sqrt{\pi}} \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1 \quad \left\{ \because \sqrt{\tfrac{1}{2}} = \sqrt{\pi} \right.$$

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

$x \in (-\infty, \infty),\ \sigma > 0$

Date __/__/__



$x = \mu$

Mean = Mode = Median

Max value $= \dfrac{1}{\sigma\sqrt{2\mu}}$

left sloped

$\beta_1 = 0$ (symetric)

$\beta_2 = 3$ (Normal distribution)

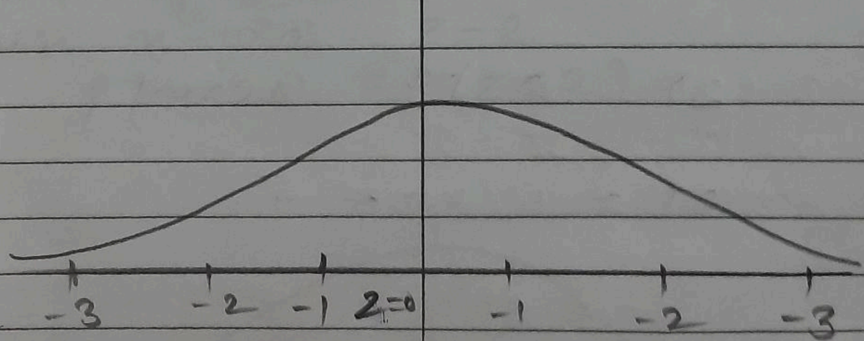
Standred normal varience / Distribution:-

$N(0, 1)$

Max varience

Probability density function:-

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty$$



$P(-1 < z < 1) = 67\%$

$P(-2 < z < 2) = 95\%$

$P(-3 < z < 3) = 99\%$

Date ___ / ___ / _____

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

$$F_X(x) = \int_{-\infty}^{x} f(x)\,dx$$

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

let $\boxed{\frac{x-\mu}{\sigma} = z} \longrightarrow$ imp

$$x = z\sigma + \mu$$

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

$$\boxed{F_X(x) = \phi(\sigma z + x)}$$

Date __ / __ / __

Standred Normal distribution

$X \sim N(0,1)$ $\qquad \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

$$\boxed{Z = \frac{x-\mu}{\sigma}}$$

Que A random variable x follow normal distribution with mean 12 and standred deviation 4. find

(i) $P(X \leq 20)$

(ii) $P(-8 < x < 20)$

(iii) $P(x \geq 12)$

(iv) $P(x \geq 20)$

Solⁿ $\mu = 12, \quad \sigma = 4$

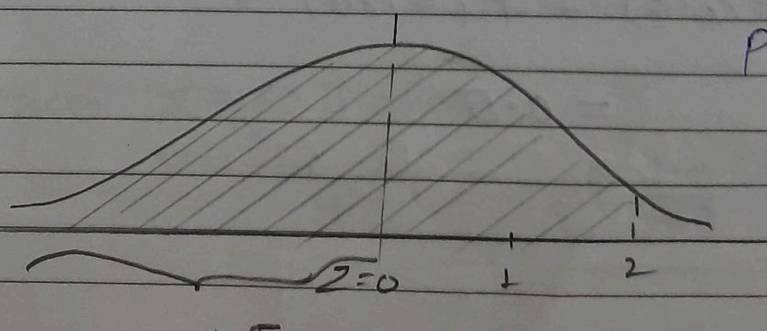
$$Z = \frac{x-\mu}{\sigma}$$

$$Z = \frac{x-12}{4}$$

(i) $P(x \leq 20)$

when $x = 20 \quad Z = 2$

$P(x \leq 20) = P(Z \leq 2)$



$P(Z \leq 2) = 0.5 + P(0 \leq Z \leq 2)$

$= 0.5 + 0.4772$

$= 0.9772$

Page No.

Date ___/___/_____

(ii) $P(-08 \leq x \leq 20)$

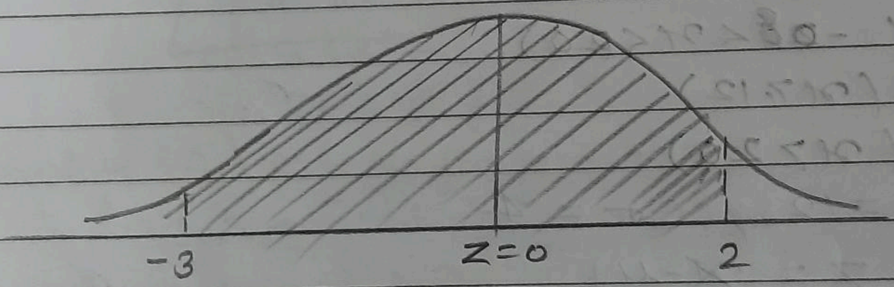
Let $x = 20$

$$Z = \frac{20-12}{4} = 2$$

$x = 08$

$$Z = \frac{08-12}{4} = -3$$

$P(0 \leq x \leq 20) = P(-3 \leq Z \leq 2)$



$P(-3 \leq Z \leq 2) = P(-3 \leq Z < 0) + P(0 \leq Z \leq 2)$

$P(0 < Z \leq 3) + P(0 \leq Z \leq 2)$

$= 0.4987 + 0.4772$

$= 0.9759$

(iii) $P(x \geq 12)$

when $x = 12$ $\quad Z = 0$

$P(x \geq 12) = P(Z \geq 0)$

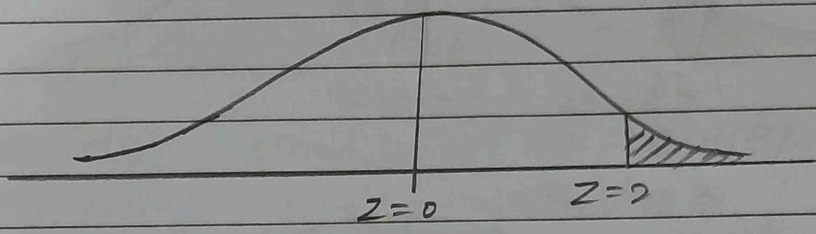
$= 0.5$

Date __/__/__

(iv) $P(x \geq 20)$

when $x = 20$, $z = 2$

$$P(x \geq 20) = P(z > 2)$$



$$P(z>2) = 0.5 - P(0<z<2)$$
$$= 0.5 - 0.4772$$
$$= 0.0228$$

Qus define normal distribution. if the hight of 300 Students are normally distributed with mean ~~$~~64.5 inches and standred deviation 3.3 inches. How many Students have hight

(i) less than 5 fit

(ii) Between 5 fit and 5.9 fit

Ans $\mu = 64.5$

$\sigma = 3.3$

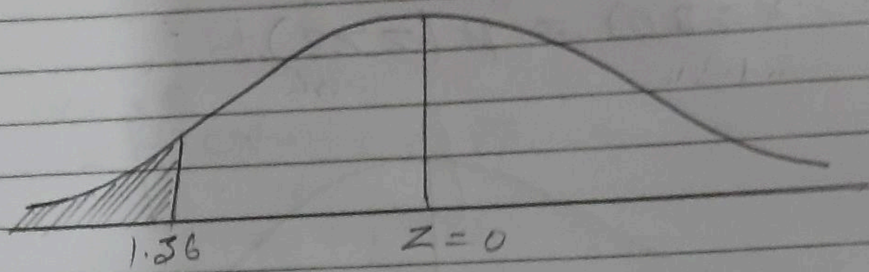
$$z = \frac{x-\mu}{\sigma} = \frac{x-64.5}{3.3}$$

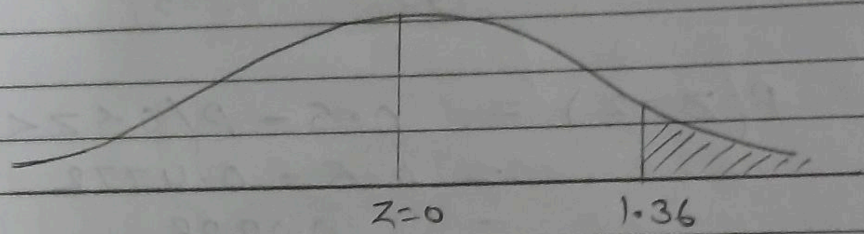
when ~~x~~ = 60 inches

$$z = \frac{60-64.5}{3.3} = \frac{-4.5}{3.3} = -1.36$$

Date __/__/____

$P(x \leq 60) = P(z < -1.36$





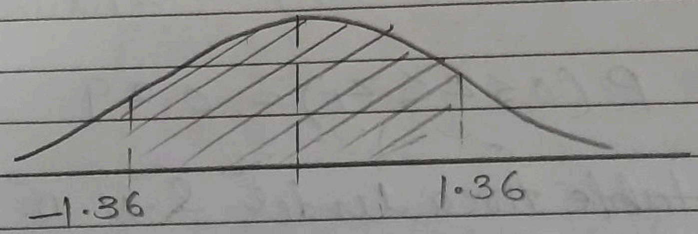
$P(Z < -1.36) = P(Z > 1.36)$
$= 0.5 - P(0 < Z < 1.36)$
$= 0.5 - 0.4131$
$= 0.0869$

No of Student $= 0.0869 \times 500$
$= 8.69 \times 3$
$= 26.07$
≈ 26 Students

(ii) $P(5 \leq X \leq 5.9)$
when $x = 69$ inches

$$Z = \frac{69 - 64.5}{3.3} = +1.36$$

Date ___/___/___



$$P(-1.36 < Z < 1.36) = 2P(Z < 1.36)$$
$$= 2 \times 0.4131$$
$$= 0.8262$$

no. of Students $= 0.8262 \times 300$
$= 82.62 \times 3$
$= 247.86$
$=$

Que A random variable X is normally distributed with mean 12 and Standard deviation 4. If $P(X \geq x_1) = 0.21$ then find the value of x_1?

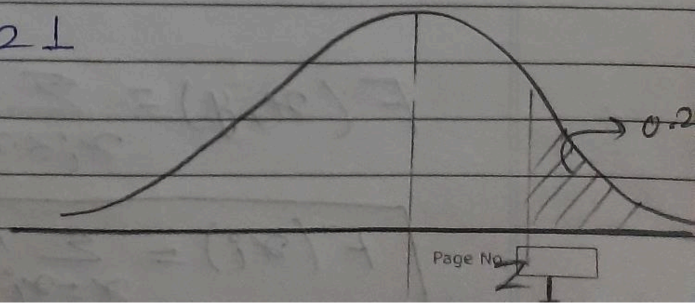
Ans $\mu = 12, \quad \sigma = 4$

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

when $x = x_1$, $Z = Z_1$

$$Z_1 = \frac{x_1 - 12}{4} \quad \text{—(i)}$$

$$P(Z \geq Z_1) = 0.21$$



Page No.

Date __/__/__

$$0.5 - P(0 \le z \le z_1) = 0.21$$

$$P(0 \le z \le z_1) = 0.29$$

From table Area under S.N.D.

$$z_1 = 0.81$$

from eqn (1)

$$0.81 = \frac{x_1 - 12}{4}$$

$$3.24 = x_1 - 12$$

$$x_1 = 12 + 2.44$$

$$= 15.24$$

#

$x \downarrow \; y \rightarrow$	y_1	y_2	...	y_n	marginal distribution of x $P(x)$
x_1	P_{11}	P_{12}	...	P_{1n}	$\sum_{j=n} P_{1j} = P_1$
x_2	P_{21}	P_{22}	...	...	P_2
⋮	⋮	⋮		⋮	⋮
x_m	P_{m1}	P_{m2}		P_{mn}	P_m
$P(y)$	q_1	q_2	...	q_n	

$$P_{ij} \geq 0 \quad \sum\sum P_{ij} = 1$$

$$F(x,y) = \sum_{x_i \le x} \sum_{y_i \le y} P(x_i, y_i)$$

$$\boxed{F(x_i) = \sum_{x = x_i} P(x)}$$

Page No.

Date ___/___/___

Continuous :-

$$f(x,y)$$

$$f(x,y) \geq 0$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

$$\boxed{F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv}$$

* If $f(x,y)$ is Probability distribution ~~of~~ function of a bivariate distribution than

$$F_x(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$F_y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

Date ___/___/___

Karl Pearson Cofficient of Correlation:-

Karl pearson gave a formula to measure the intensity or degree of linear relationship between two variables, known as Correlation Coefficient $r(x,y)$ as

$$r(x,y) = r_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y} \quad \text{—— (1)}$$

$$Cov(x,y) = \mu_{11} = E[(x-E(x))(y-E(y))]$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

$$\sigma_x^2 = E[(x-E(x))^2] = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

and

$$\sigma_y^2 = E[(y-E(y))^2] = \frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2$$

from (1)

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

or

$$r_{xy} = \frac{\frac{1}{n}\sum_{i=1}^{n}x_i y_i - \bar{x}\bar{y}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}y_i^2-\bar{y}^2}}$$

Date ___/___/______

Properties of Correlation Coefficient – r_{xy}:

(i) Correlation Coefficient is independent of Change of origin and scale.

let

$$U = \frac{X-a}{h} \quad \text{and} \quad V = \frac{Y-b}{k}$$

$$r_{xy} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{Cov(U,V)}{\sigma_U \sigma_V}$$

(ii) The Coefficient of Correlation lies between -1 and 1.

i.e. $-1 \leq r_{xy} \leq 1$.

(iii) If $r = \pm 1$, a perfect correlation exists between the variables and if $r = 0$ the variable are uncorrelated.

(iv) Two independent - variables are uncorrelated but Converse is not true.

X	-3	-2	-1	0	1	2	3
Y	9	4	1	0	1	4	9
XY	-27	-8	-1	0	1	8	27

$\Sigma X = 0$, $\Sigma Y = 28$ and $\Sigma XY = 0$

Date __/__/____

Now $\bar{X} = \frac{\Sigma x}{n} = 0$ and $Cov(x,y) = \frac{1}{n}\Sigma xy - \bar{x}\bar{y} = 0$

$\Rightarrow r = \frac{Cov(x,y)}{\sigma_x \sigma_y} = 0$

Ex. 1. Calculate the Correlation Coefficiant for the following heights (in inches) of fathers (X) and their sons (Y):

X : 65 66 67 67 68 69 70 72

Y : 67 68 65 68 72 72 69 71

Sol^n Let $U = X - 68$ and $V = Y - 69$

X	Y	$U = X-68$	$V = Y-69$	U^2	V^2	UV
65	67	-3	-2	9	4	6
66	68	-2	-1	4	1	2
67	65	-1	-4	1	16	4
67	68	-1	-1	1	1	1
68	72	0	3	0	9	0
69	72	1	3	1	9	3
70	69	2	0	4	0	0
72	71	4	2	16	4	8
Total		0	0	36	44	24

Now $\bar{u} = \frac{1}{n}\Sigma u = 0$, $\bar{v} = \frac{1}{n}\Sigma v = 0$

$Cov(u,v) = \frac{1}{n}\Sigma uv - \bar{u}\bar{v} = \frac{1}{8} \times 24 = 3$

Date __/__/__

$\sigma_u = \sqrt{\frac{1}{n}\Sigma u^2 - \bar{u}^2} = \sqrt{\frac{36}{8}} = \sqrt{4.5} = 2.121$

$\sigma_v = \sqrt{\frac{1}{n}\Sigma v^2 - \bar{v}^2} = \sqrt{\frac{1}{8} \times 44} = \sqrt{5.5} = 2.345$

Hence

$\gamma = \frac{Cov(x,y)}{\sigma_x \sigma_y} = \frac{Cov(u,v)}{\sigma_u \sigma_v} = \frac{3}{2.121 \times 2.345} = 0.6032$

Exam. 2. Calculate the karl Pearson's cofficient of correlation of the following data :

x	25	27	30	35	33	28	36
y	19	22	27	28	30	23	28

Solⁿ Here $\bar{x} = \frac{214}{7} = 30.5$ and $\bar{y} = \frac{177}{7} = 25.2$

Let $U = X - 31$ and $V = Y - 25$

X	Y	U	V	U^2	V^2	UV
25	19	-6	-6	36	36	36
27	22	-4	-3	16	9	12
30	27	-1	2	1	4	-2
35	28	4	3	16	9	12
33	30	2	5	4	25	10
28	23	-3	-2	9	4	6
36	28	5	3	25	9	15
Total		-3	2	107	96	89

Date ___/___/_____

Now $\bar{u} = \frac{1}{n}\Sigma u = \frac{-3}{7} = -0.429$,

$$\bar{v} = \frac{2}{7} = 0.286$$

$$Cov(u,v) = \frac{1}{n}\Sigma uv - \bar{u}\bar{v} = \frac{1}{7}(89) - (-0.429)(0.286)$$
$$= 12.592$$

$$\sigma_u = \sqrt{\frac{1}{n}u^2 - \bar{u}^2} = \sqrt{15.1017} = 3.886$$

$$\sigma_v = \sqrt{\frac{1}{n}\Sigma v^2 - \bar{v}^2} = \sqrt{13.6325} = 3.692$$

Hence $r = \frac{Cov(u,v)}{\sigma_u \sigma_v} = \frac{12.592}{3.886 \times 3.692} = 0.878$

Rank Correlation :-

Let us suppose that for a group of n individuals grades or ranks (x_i, y_i) $i = 1, 2, 3, \ldots n$ are given with respect to two characteristics A and B respectively. Then Spearman's rank correlation coefficient for non repeated rank is

$$\rho = \frac{1 - 6\Sigma d_i^2}{n(n^2-1)} \quad \text{where } d_i = x_i - y_i$$

** $-1 \leq \rho \leq 1$

Rank correlation factor for repeated Ranks

for repeated ranks, a correlation factor is required in the formula. If m is the number of times a rank is repeated then the factor $\frac{m(m^2-1)}{12}$ is to be added to Σd_i^2. This correlation factor is added for each repeated rank.

$$\rho = 1 - \frac{6\,[\Sigma d_i^2 + C.f.]}{n(n^2-1)}$$

Ex.1. The ranking of ten students in two subjects A and B are as follows.

A	3	5	8	4	7	10	2	1	6	9
B	6	4	9	8	1	2	3	10	5	7

Soln

A	B	d_i	d_i^2
3	6	-3	9
5	4	1	1
8	9	-1	1
4	8	-4	16
7	1	6	36
10	2	8	64
2	3	-1	1
1	10	-9	81
6	5	1	1
9	7	2	4
Total		0	214

Page No.

Spearman rank correlation coefficient:

$$r = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 214}{10(10^2-1)} = 1 - \frac{6 \times 214}{10 \times 99} = 1 - 1.3 = -0.3$$

Ex. 2. Calculate the Coefficient of Correlation from the data given below by the method of differences:

x :	78	89	97	69	59	79	68	57
y :	125	137	156	112	107	136	123	108

Sol^n

x	y	Rank in x x	Rank in y y	Rank diff d_i	d_i^2
78	125	4	4	0	0
89	137	2	2	0	0
97	156	1	1	0	0
69	112	5	6	-1	1
59	107	7	8	-1	1
79	136	3	3	0	0
68	123	6	5	1	1
57	108	8	7	1	1
Total				$\Sigma d_i = 0$	$\Sigma d_i^2 = 4$

$$r = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 4}{8(64-1)}$$

Date __/__/__

$$= 1 - \frac{3}{63}$$

$$= \frac{20}{21} = 0.95$$

Ex. 3. In the following table are recorded data showing the test scores made by 10 salesman on an intelligence test and their weekly sales :-

~~Boss.~~ Salesmen	1	2	3	4	5	6	7	8	9	10
Test scores	50	70	50	60	80	50	90	50	60	60
Sales('000Rs)	25	60	45	50	45	20	55	30	45	30

Calculate the rank correlation coefficients between intelligence and efficiency in salesmanship.

Solⁿ

Salesman	Intelligence Testscore	Rank (x)	Sales Amount (000 Rs)	Rank (y)	$d = x - y$	d^2
1	50	8.5	25	9	−0.5	0.25
2	70	3	60	1	2	4
3	50	8.5	45	5	3.5	12.25
4	60	5	50	3	2	4
5	80	2	45	5	−3	9
6	50	8.5	20	10	−1.5	2.25
7	90	1	55	2	−1	1
8	50	8.5	30	7.5	1	1
9	60	5	45	5	0	0
10	60	5	30	7.5	−2.5	6.25
	Total				0	40

Date __/__/__

In the two series together, there are in all 4 ties, viz. one tie involving 4 individuals, 2 ties involving 3 individuals each and 1 tie with 2 individuals

Hence

$$C.F. = \frac{\Sigma m(m^2-1)}{12} = \frac{4^3-4}{12} + \frac{3^3-3}{12} + \frac{3^3-3}{12} + \frac{2^3-2}{12}$$

$$= 5 + 2 + 2 + 0.5$$

$$= 9.5$$

$$\rho = 1 - \frac{6[\Sigma d^2 + C.F.]}{n(n^2-1)}$$

$$= 1 - \frac{6[40 + 9.5]}{10(10^2-1)}$$

$$= 0.70$$

Ex.4. Calculate the Coefficient of Correlation from the following data:

X: 45, 56, 39, 54, 45, 40, 56, 60, 30, 36
Y: 40, 36, 30, 44, 36, 32, 45, 42, 20, 36

Sol

X	Rank of X (x)	Y	Rank of Y (y)	$d = x - y$	d^2
45	5.5	40	4	1.5	2.25
56	2.5	36	6	-3.5	12.25
39	8	30	9	-1	1
54	4	44	2	2	4

Date ___/___/___

45	5.5	36	6	-0.5	0.25
40	7	32	8	-1	1
56	2.5	45	1	1.5	2.25
60	1	42	3	-2	4
30	10	20	10	0	0
36	9	36	6	3	9
				$\Sigma d = 0$	$\Sigma d^2 = 36$

In the two series together, there are in all 3 ties, viz. 2 ties involving 2 inviduals and 1 tie with 3 inviduals. Hence

$$C.F. = \frac{1}{12}\{(2^3-2)+(2^3-2)+(3^3-3)\}$$

$$= \frac{36}{12} = 3$$

and $$\rho = 1 - \frac{6[\Sigma d^2 + C.F.]}{n(n^2-1)}$$

$$= 1 - \frac{6[36+3]}{10(10^2-1)}$$

$$= 1 - \frac{6 \times 39}{990}$$

$$= 1 - 0.2364$$

$$= 0.7636$$

Date __/__/__

REGRESSION :-

The word 'regression' is used to denote estimation or prediction of the average value of one variable for a specified value of the other variable. The estimation is done by means of suitable equations, derived on the basis of available bivariate data: such an equation is known as a regression equation and its geometrical representation is called a Regression curve.

In linear regression (or simple regression) the relationship between the variable is assumed to be linear. The estimate of y (say Y) is obtained from an equation of the form

$$(Y - \bar{y}) = b_{yx}(X - \bar{x}) \quad \text{—— (1)}$$

where

$$b_{yx} = \frac{Cov(x,y)}{\sigma_x^2} = r\frac{\sigma_y}{\sigma_x}$$

and the estimation of x (say X) from another equation of the form

$$(X - \bar{x}) = b_{xy}(Y - \bar{y})$$

where $b_{xy} = \dfrac{Cov(x,y)}{\sigma_y^2} = r\dfrac{\sigma_x}{\sigma_y}$

Equation (1) is known as Regression equation of y on x and equation (2) as Regression equation of x on y.

** The Coefficient b_{yx} is known as the Regression coefficient of y on x.

** The Coef. b_{xy} is called the Regression Coefficient of x on y.

The geometrical representation of linear regression equations (1) and (2) are known as Regression lines.
These lines are 'best fitting' Straight lines obtained by the method of least squares.

Properties of linear Regression :-

(1) There are two linear regression equations
(i) Regression equation of y on x

$$(y-\bar{y}) = b_{yx}(x-\bar{x})$$

or $(y-\bar{y}) = r\frac{\sigma_y}{\sigma_x}(x-\bar{x})$

or $\frac{(y-\bar{y})}{\sigma_y} = \frac{r(x-\bar{x})}{\sigma_x}$ where r is karl person's Coeff. of correlation.

Date __ / __ / ____

(ii) Regression equation of x on y

$$(x-\bar{x}) = b_{xy}(y-\bar{y})$$

or $(x-\bar{x}) = r\dfrac{\sigma x}{\sigma y}(y-\bar{y})$

or $\left(\dfrac{x-\bar{x}}{\sigma x}\right) = r\left(\dfrac{y-\bar{y}}{\sigma y}\right)$ where r is karl pearson's coeff. of correlation.

(2) The product of two regression coefficient is equal to the square of correlation coefficient

$$b_{yx} \cdot b_{xy} = r^2$$

(3) r, b_{yx} and b_{xy} all have the same sign. If the correlation coefficient r is zero, the regression coefficient b_{yx} and b_{xy} are also zero.

(4) The regression lines always intersect at the point $(\bar{x}, \bar{y})$. The slopes of the regression line of y on x and the regression line of x on y are respectively b_{yx} and $\dfrac{1}{b_{xy}}$.