Computer Science & DA

Probability and Statistics

Continuous Random variable

Lecture No. 05

By- Dr. Puneet Sharma Sir

Recap of previous lecture

Topic Normal Distribution 2 CLT, NST, CDF.

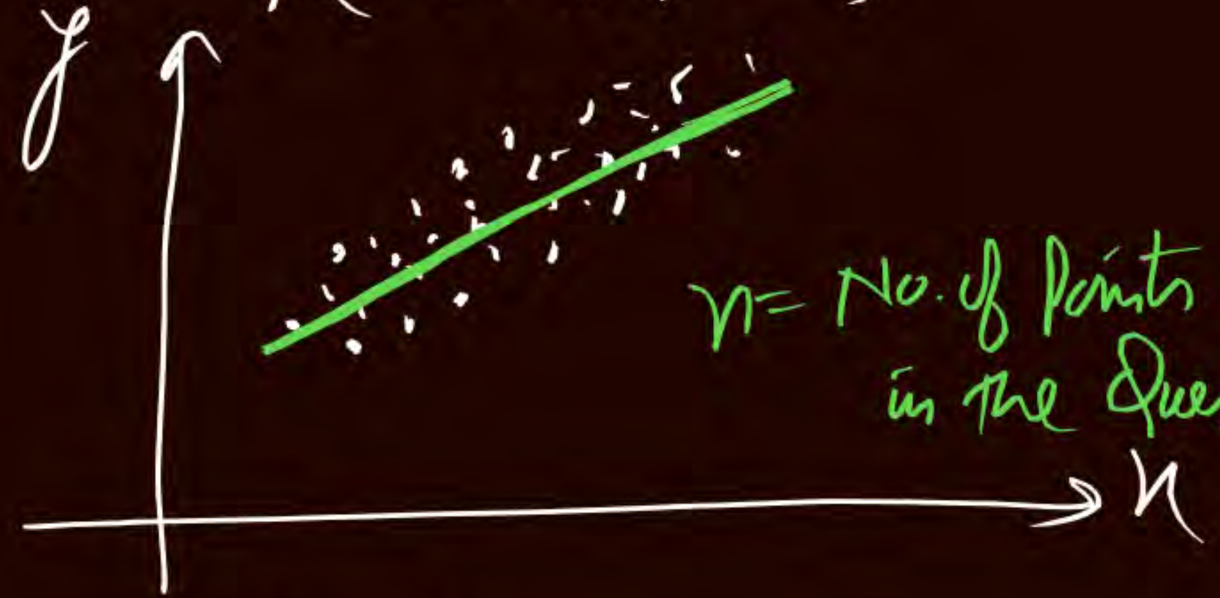
Topics to be Covered

Topic $\underbrace{\text{Bivariate}}_{x,\,y}$ Random variable (Basics)

Curve fitting

$x=$	x_1	x_2	x_3	- - - - - - -	x_n
$y=$	y_1	y_2	y_3	- - - - - - - -	y_n

$y(25\text{ mtrs}) = ?$



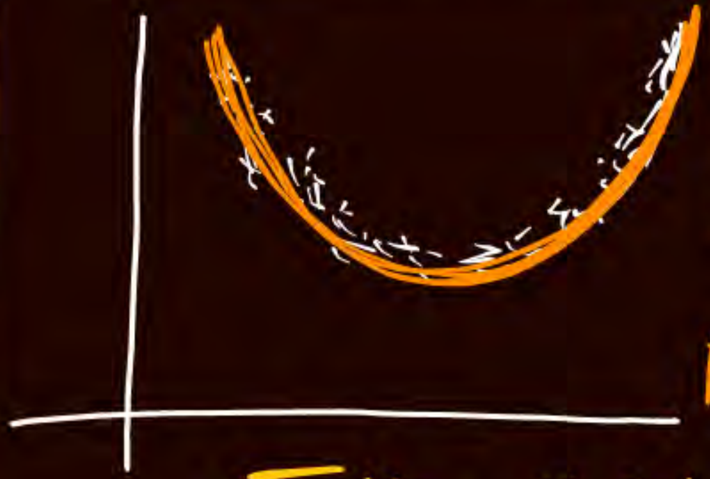
$n =$ No. of Points given in the Question

Let the line of Best fit is $\boxed{y = a + bx}$ — (1)

To find a & b we will solve following equ^n's:

$$\left.\begin{aligned} \sum y &= na + b\sum x \\ \sum xy &= a\sum x + b\sum x^2 \end{aligned}\right\}$$ Normal equ^n for line of Best fit.

(*) Parabola of Best fit — is given as;



$\boxed{y = a + bx + cx^2}$ — (1)

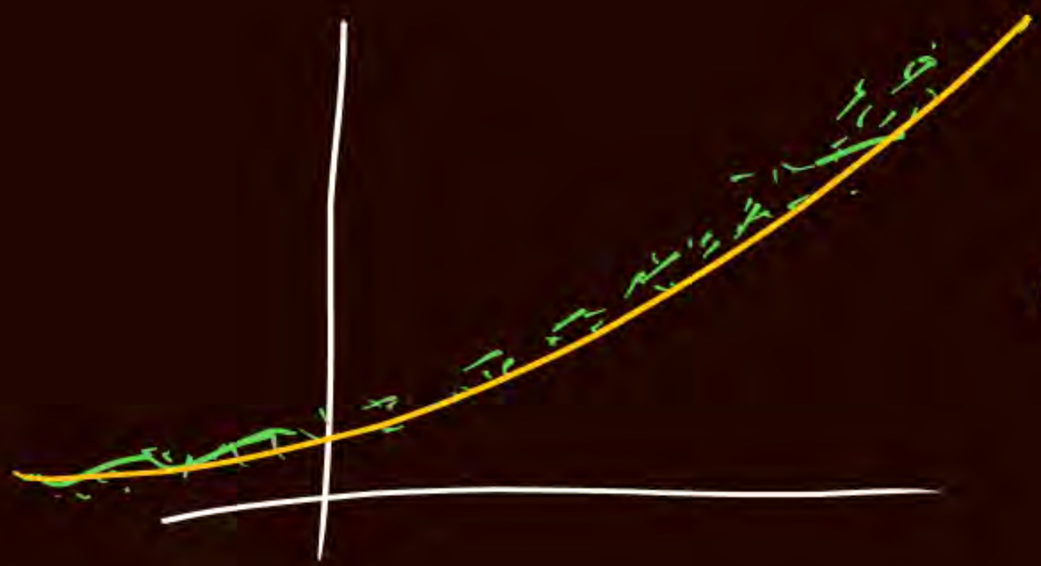
Now we will solve N equ^n's;

$$\sum y = na + b\sum x + c\sum x^2$$

$$\sum xy = a\sum x + b\sum x^2 + c\sum x^3$$

$$\sum x^2y = a\sum x^2 + b\sum x^3 + c\sum x^4$$

By solving these N equ^n's, we can find a, b, c & Hence Parabola of Best fit also.



$$y = ae^{bx}$$

Q. find the line of Best fit for the following Data

$x=$	1	2	3	4	5
$y=$	3	5	7	9	11

M-I By observation; $y = 1 + 2x$

this is Called Line of Best fit for given data.

M-II

x	y	x^2	xy
1	3	1	3
2	5	4	10
3	7	9	21
4	9	16	36
5	11	25	55
Σ 15	35	55	125

N. equ^n are;

$$\left.\begin{aligned} \Sigma y &= na + b\Sigma x \\ \Sigma xy &= a\Sigma x + b\Sigma x^2 \end{aligned}\right\} \Rightarrow \begin{aligned} 35 &= 5a + 15b \\ 125 &= 15a + 55b \end{aligned}$$

$a = 1,\ b = 2$

So line of Best fit is $y = a + bx$

$$\boxed{y = 1 + 2x}$$

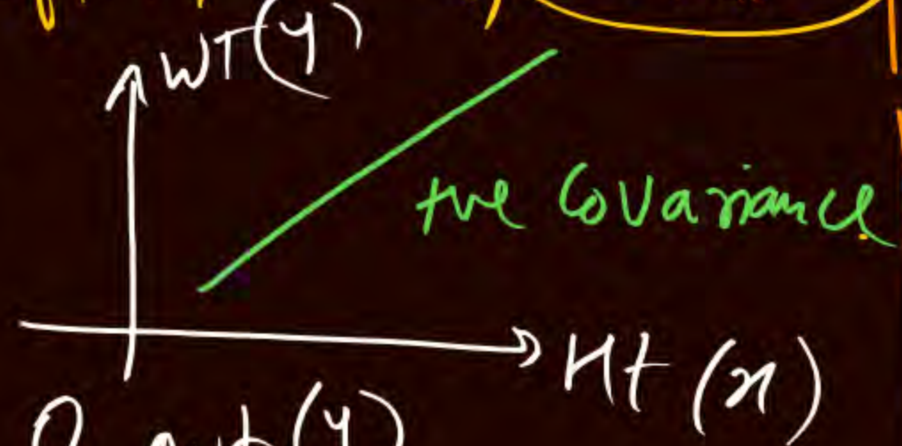
Correlation and Regression

Covariance

1. $\boxed{Cov(X,Y) = E(XY) - E(X).E(Y)}$

2. if X & Y are Ind, $Cov(X,Y)=0$

eg



$-\infty < Cov < \infty$

it measures only direction

Correlation

1. Correlation coeff $\boxed{r = \dfrac{Cov(X,Y)}{\sigma_x \sigma_y}}$ Karl-Pearson

2. If x & y are Ind, $r = 0$

3. $-1 \le r \le 1$

(i) if $r = 1$ then X & Y are perfectly Correlated in +ve sense

(ii) if $r = -1$ " " " " -ve "

(*) It measures the direction as well as strength of relationship

REGRESSION

if supply is constant then

$\overset{(y)}{\text{Demand}} \propto \frac{1}{\text{Price}} \Rightarrow y \propto \frac{1}{x}$ (customer)

Price $\propto$ Demand $\Rightarrow x \propto y$ (shopkeeper)

Let Demand $= y$
Price $= x$

$y = f(x) \Rightarrow W = f(H)$
or
$x = f(y) \Rightarrow H = f(W)$

R. line Y on X i.e. $Y - \bar{Y} = b_{YX}(X - \bar{X})$ where $b_{YX} = r\frac{\sigma_Y}{\sigma_X}$

R. line X on Y i.e. $X - \bar{X} = b_{XY}(Y - \bar{Y})$ & $b_{XY} = r\frac{\sigma_X}{\sigma_Y}$

Here b_{YX} & b_{XY} are called R. Coeff.

Note → ① For R-Coeff we have following possibilities;

(i) Either both b_{yx} & b_{xy} are less than one

(ii) if $\underbrace{\text{one is greater than one}}$ then $\underbrace{\text{another must be less than one}}$

$$\text{if } b_{yx} > 1 \;\underset{\not\Leftarrow}{\Longrightarrow}\; b_{xy} < 1$$

OR

$$\text{if } b_{xy} > 1 \;\underset{\not\Leftarrow}{\Longrightarrow}\; b_{yx} < 1$$

(iii)

$$\left.\begin{aligned} b_{yx} &= r\frac{\sigma_y}{\sigma_x} \\ \& \; b_{xy} &= r\frac{\sigma_x}{\sigma_y} \end{aligned}\right\} \Rightarrow \begin{aligned} b_{yx}\cdot b_{xy} &= \left(r\frac{\sigma_y}{\sigma_x}\right)\left(r\frac{\sigma_x}{\sigma_y}\right) \\ b_{yx}\, b_{xy} &= r^2 \end{aligned}$$

i.e. $\boxed{r = \pm\sqrt{b_{yx}\cdot b_{xy}}}$

i.e. r is the GM of R-Coeff

(a) AM (b) GM (c) HM (d) None

Reason for Point (ii) →

$\because -1 \le r \le 1$ & r is GM Hence proved

(iv) r, b_{yx} & b_{xy} have Same sign

(v) Intersecting Point of two R-lines is nothing but $(\bar{x}, \bar{y})$

eg GM of 4 & 9 is = ? $= \sqrt{4\times 9} = \sqrt{36} = +6$

eg GM of -4 & -9 is = ? $= \cancel{\sqrt{(-4)(-9)} = \sqrt{36} = +6}$

$-9 \quad -6 \quad -4$

$= -\sqrt{(-4)(-9)} = -\sqrt{36} = -6$

eg GM of -9 & 4 = ? = N.D

(*) GM of two +ve Nos a & b $= +\sqrt{ab}$

& " " -ve " " " $= -\sqrt{ab}$

if one is +ve & other is -ve then GM = N.D

(VII) Angle b/w two R-lines is ?

$$\boxed{\tan\theta = \left(\frac{1-r^2}{r^2}\right)\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}}$$

eg: Two R-lines will be $\perp^r$ if r = ?

(a) 1 (b) -1 (c) 0 (d) ∞

At $\theta = 90^\circ$, $\tan\theta = \infty$

$\Rightarrow r = 0$

Correlation

① It measures direction and strength of relationship

② it is not Based on Cause and Effect ie we can't predict Effect due to Cause

③ Sometimes we may see senseless Correlation

Regression

① it measures the Variation of one variable w.r.to other Variable

② It is useful to predict Cause & Effect in the relationship.

$y = f(x)$ or $x = f(y)$

Effect = f(Cause) Effect = f(Cause)

③ Here we will never see senseless relation.

Q: For the Data
x = 1, 2, 3, 4, 5
y = 3, 5, 7, 9, 11

x	1	2	3	4	5
y	3	5	7	9	11

evaluate Correlation Coeff, R. Coeff, R. line Y on X and R. line X on Y ? $n = 5$

Sol:

x	y	xy	x^2	y^2
1	3	3	1	9
2	5	10	4	25
3	7	21	9	49
4	9	36	16	81
5	11	55	25	121
Σ 15	35	125	55	285

$E(X) = \frac{\Sigma X}{n} = \frac{15}{5} = 3$, $E(X^2) = \frac{\Sigma X^2}{n} = \frac{55}{5} = 11$

$E(Y) = \frac{\Sigma Y}{n} = \frac{35}{5} = 7$, $E(Y^2) = \frac{\Sigma Y^2}{n} = \frac{285}{5} = 57$

$E(XY) = \frac{\Sigma XY}{n} = \frac{125}{5} = 25$

$Var(X) = E(X^2) - E^2(X) = 11 - (3)^2 = 2 \Rightarrow \sigma_x = \sqrt{2}$

$Var(Y) = E(Y^2) - E^2(Y) = 57 - (7)^2 = 8 \Rightarrow \sigma_y = 2\sqrt{2}$

$Cov(X,Y) = \boxed{E(XY) - E(X).E(Y)} = 25 - (3)(7) = 4$

(1) $r = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{4}{(\sqrt{2})(2\sqrt{2})} = 1$ (Perfect Correlation)

(2) $b_{yx} = r.\frac{\sigma_y}{\sigma_x} = 1.\frac{2\sqrt{2}}{\sqrt{2}} = 2$

$b_{xy} = r\frac{\sigma_x}{\sigma_y} = 1.\frac{\sqrt{2}}{2\sqrt{2}} = \frac{1}{2}$

③ R line Y on X :→

$$Y - \bar{Y} = b_{YX}(X - \bar{X})$$

$$y - 7 = 2(x - 3)$$

$$\boxed{y = 2x + 1} \longrightarrow (1)$$

④ R line X on Y :→

$$X - \bar{X} = b_{XY}(Y - \bar{Y})$$

$$x - 3 = \frac{1}{2}(y - 7)$$

$$\boxed{x = \frac{y}{2} - \frac{1}{2}} \longrightarrow (2)$$

(*) $\because r = 1$ that's why eqn (2) can be calculated by using (1)

But we should avoid this process

(*) Intersecting Point is $(\bar{x}, \bar{y}) = (3, 7)$

Q. If $\bar{x}=10$, $\bar{y}=90$, $\sigma_x=3$, $\sigma_y=12$, $r=0.8$ then find R.line y on x and x on y

Sol: $b_{yx} = r\dfrac{\sigma_y}{\sigma_x} = 0.8\left(\dfrac{12}{3}\right) = 3.2$

R.line y on x → $y-\bar{y} = b_{yx}(x-\bar{x})$

$y-90 = 3.2(x-10)$

$\boxed{y = 3.2x + 58}$ Ans

$b_{xy} = r\dfrac{\sigma_x}{\sigma_y} = 0.8\left(\dfrac{3}{12}\right) = 0.2$

R.line x on y → $x-\bar{x} = b_{xy}(y-\bar{y})$

$x-10 = 0.2(y-90)$

$\boxed{x = 0.2y - 8}$ Ans

(*) Intersecting Point $= (\bar{x},\bar{y}) = (10,90)$

ESE: Find the intersecting Point of two R.lines? $=(10,90)$

Q.

1	2	3	4	5	6	7
9	8	10	12	11	13	14

Find $\underbrace{\text{Correlation Coeff}}_{r=?}$, $\underbrace{\text{R. Coeff}}_{b_{yx}\ \&\ b_{xy}}$ & $\underbrace{\text{R. line } y \text{ on } x}$ and $\underbrace{\text{R. Line } x \text{ on } y}$?

HW $r = b_{yx} = b_{xy} = 0.929$

R. line y on x : $y = 0.929x + 7.824$

R line x on y : $x = 0.929y - 6.219$

Double Integral Working Rule →

Case I: If Both the variables have constant limits and integrand is an Explicit funcⁿ of x & y we can integrate separately.

Case II: → If Both the variables have constant limits and integrand is an Implicit funcⁿ of x & y. then we can integrate (one by one) in any order

Case III: (for vertical strip) → if y has variable limits then we should 1st integrate dy keeping x constant.

Case IV: → (for Horizontal strip) → if x has " " " " " " " dx keeping y constant.

Case V: If Both have variable limits → Semselen Q.

Note → the concept of Explicit funcⁿ is applicable only when both the limits are const.

Q. $I = \int_0^3 \int_0^2 x^2 y \, \underline{dx\,dy} = ? = \int_{y=0}^{3} \int_{x=0}^{2} x^2 y \, dx\,dy = \int_{x=0}^{2} x^2 dx \times \int_{y=0}^{3} y\,dy = \left(\frac{x^3}{3}\right)_0^2 \left(\frac{y^2}{2}\right)_0^3$

Explicit

$= \frac{8}{3} \times \frac{9}{2} = 4 \times 3 = 12$ //

Q. $\int_0^1 \int_0^2 e^{x+y} \, \underline{dy\,dx} = ? = \int_{x=0}^{1} \int_{y=0}^{2} e^{x+y} \, dy\,dx = \int_{x=0}^{1} e^x dx \cdot \int_{y=0}^{2} e^y dy = (e-1)(e^2-1)$ //

Explicit

Q. $\int_0^{\pi/2} \int_0^{\pi/2} \sin(x+y) \, \underline{dx\,dy} = ? = \int_{y=0}^{\pi/2} \int_{x=0}^{\pi/2} \sin(x+y) \, dx\,dy$ → M-I, M-II

Implicit

(M-I) $I = \int_{y=0}^{\pi/2} \left[\int_{x=0}^{\pi/2} \sin(x+y)\, dx \right] dy = \int_{y=0}^{\pi/2} \left\{ -\cos(x+y) \right\}_{x=0}^{\pi/2} . dy = \int_{y=0}^{\pi/2} - \left\{ \cos\left(\frac{\pi}{2}+y\right) - \cos y \right\} dy$

$= \int_{y=0}^{\pi/2} - \left\{ -\sin y - \cos y \right\} dy = \int_{0}^{\pi/2} (\sin y + \cos y)\, dy = \left(-\cos y + \sin y \right)_{y=0}^{\pi/2} = \cdots\cdots = 2$

(M-II) $I = \int_{x=0}^{\pi/2} \left[\int_{y=0}^{\pi/2} \sin(x+y)\, dy \right] . dx = \int_{x=0}^{\pi/2} \left\{ -\cos(x+y) \right\}_{y=0}^{\pi/2} dx = \int_{x=0}^{\pi/2} - \left\{ \cos\left(\frac{\pi}{2}+x\right) - \cos x \right\} dx$

$= \int_{x=0}^{\pi/2} \left\{ \sin x + \cos x \right\} dx = \left(-\cos x + \sin x \right)_{0}^{\pi/2} = \cdots\cdots = 2$

Q. $I = \int_0^2 \int_0^{x/2} e^{x^2}\, dy\, dx = ? = \int_{x=0}^{2}\int_{y=0}^{x/2} e^{x^2} dy\, dx = \int_{x=0}^{2}\left[\int_{y=0}^{x/2} 1.dy\right](e^{x^2}dx)$

Case III

$= \int_{x=0}^{2}\left(\frac{x}{2}\right) . e^{x^2} dx$ ═══ $\int_0^4 \frac{1}{2} e^t . \frac{dt}{2} = \frac{1}{4}\left(e^t\right)_0^4 = \frac{e^4-1}{4}$

put $x^2 = t$ → at $x=0$, $t=0$; at $x=2$, $t=4$

$x\,dx = \frac{dt}{2}$

Q. $I = \int_0^1 \int_{y^2}^{5y} (1)\, dx\, dy = ? = \int_{y=0}^{1}\int_{x=y^2}^{5y} (1)\, dx\, dy = \int_{y=0}^{1}\left[\int_{x=y^2}^{5y} (1)\, dx\right] dy = \int_{y=0}^{1} (5y - y^2)\, dy$

Case IV

$= \left(\frac{y^{3/2}}{3/2} - \frac{y^3}{3}\right)_0^1 = \cdots = \frac{1}{3}$

Q. $I = \int_0^2 \int_0^{x^2} (xy^2)\,dx\,dy = \,? = \int_{x=0}^{2}\int_{y=0}^{x^2} (xy^2)\,dy\,dx$ → M-I → ; M-II →

M-I wrong Method →

$$I = \int_{x=0}^{2} x\,dx \times \int_{y=0}^{x^2} y^2\,dy$$

$$= \left(\frac{x^2}{2}\right)_0^2 \left(\frac{y^3}{3}\right)_0^{x^2}$$

$$= (2)\left(\frac{x^6}{3}\right) = \frac{2}{3}x^6$$

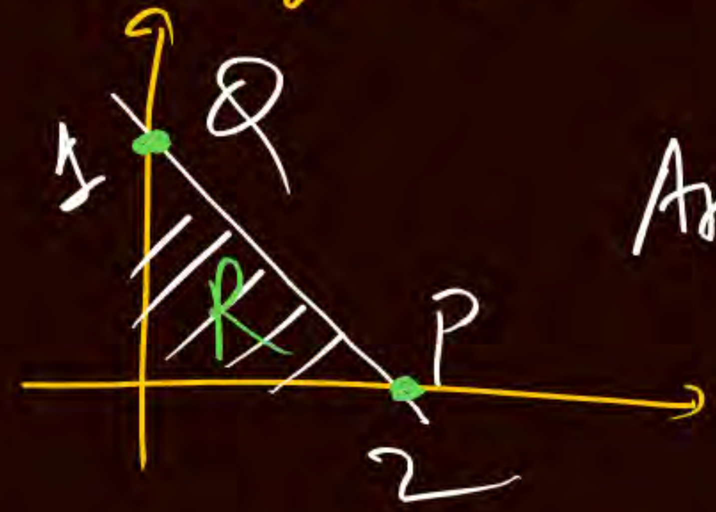
M-II Correct App → (it is Based on Case III) →

$$I = \int_{x=0}^{2}\left[\int_{y=0}^{x^2} y^2\,dy\right](x\,dx) = \int_{x=0}^{2}\left(\frac{x^6}{3}\right)x\,dx$$

$$= \frac{1}{3}\int_{x=0}^{2} x^7\,dx = \frac{1}{3}\left(\frac{x^8}{8}\right)_0^2 = \frac{2^8}{3\times 8} = \frac{2^5}{3} = \frac{32}{3}$$

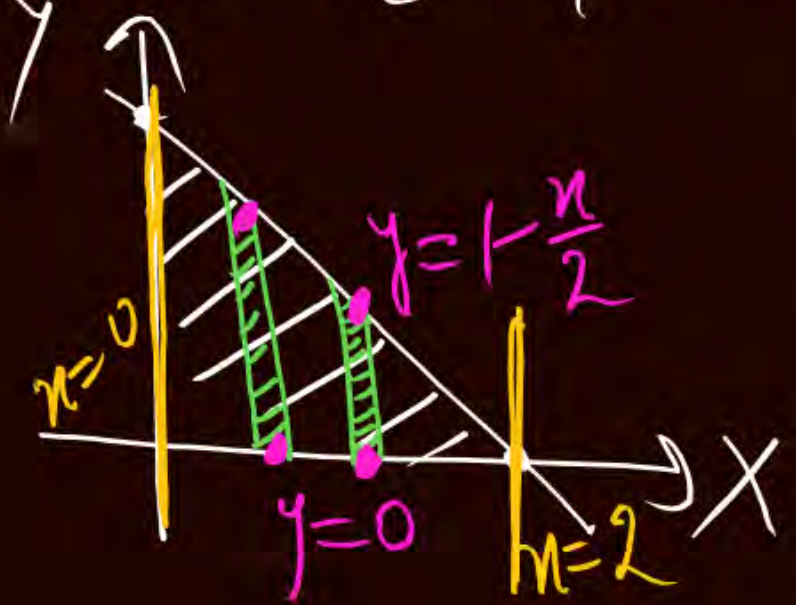
Q. Evaluate $\iint_R xy\,dx\,dy = ?$

over the Region R shown in the diag?



$$\text{Area} = \frac{1}{2}(2)(1) = 1$$

Sol: PQ: $\frac{x}{2} + \frac{y}{1} = 1 \Rightarrow y = \left(1 - \frac{x}{2}\right)$



$0 \le y \le 1 - \frac{x}{2}$

$0 \le x \le 2$

$$I = \int_{x=0}^{2}\int_{y=0}^{1-\frac{x}{2}} xy\,dy\,dx = \int_{x=0}^{2}\left[\int_{y=0}^{1-\frac{x}{2}} y\,dy\right]x\,dx$$

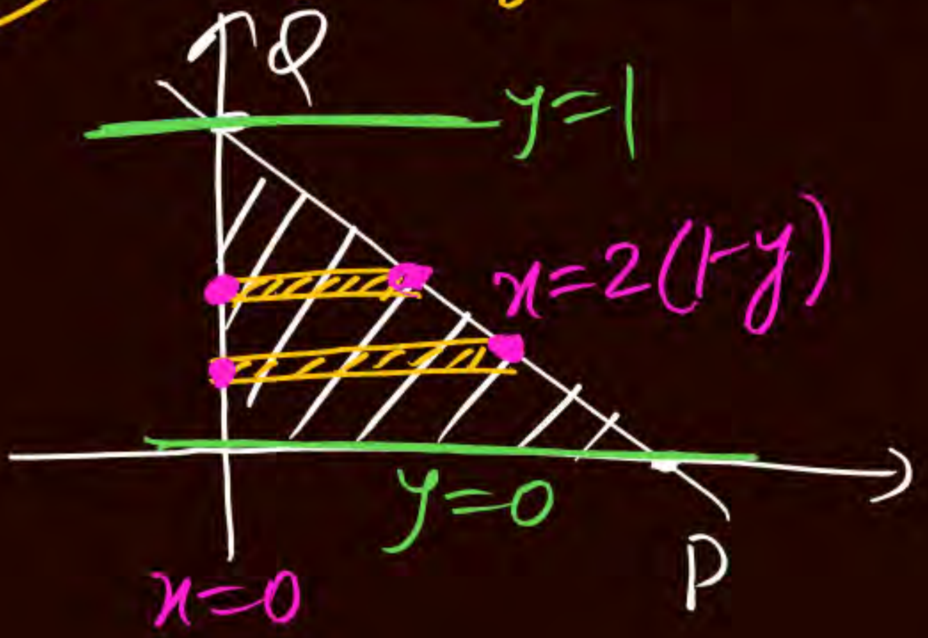
$$= \int_{x=0}^{2}\left[\frac{\left(1-\frac{x}{2}\right)^2}{2}\right](x\,dx) = \frac{1}{8}\int_0^2 (4 + x^2 - 4x)\,x\,dx$$

$$= \frac{1}{8}\int_0^2 (4x + x^3 - 4x^2)\,dx$$

$$= \frac{1}{8}\left[2x^2 + \frac{x^4}{4} - \frac{4x^3}{3}\right]_0^2$$

$$= \cdots \qquad \text{Ans} = \frac{1}{6}$$

M-II Using Horizontal Strip :-



$PQ: \frac{x}{2} + \frac{y}{1} = 1 \Rightarrow x = 2(1-y)$

ie $0 \le x \le 2(1-y)$

$0 \le y \le 1$

$$I = \iint_R xy\,dx\,dy = \int_{y=0}^{1} \int_{x=0}^{2(1-y)} xy\,dx\,dy$$

$$= \int_{y=0}^{1} \left[\int_{x=0}^{2(1-y)} x\,dx \right] y\,dy$$

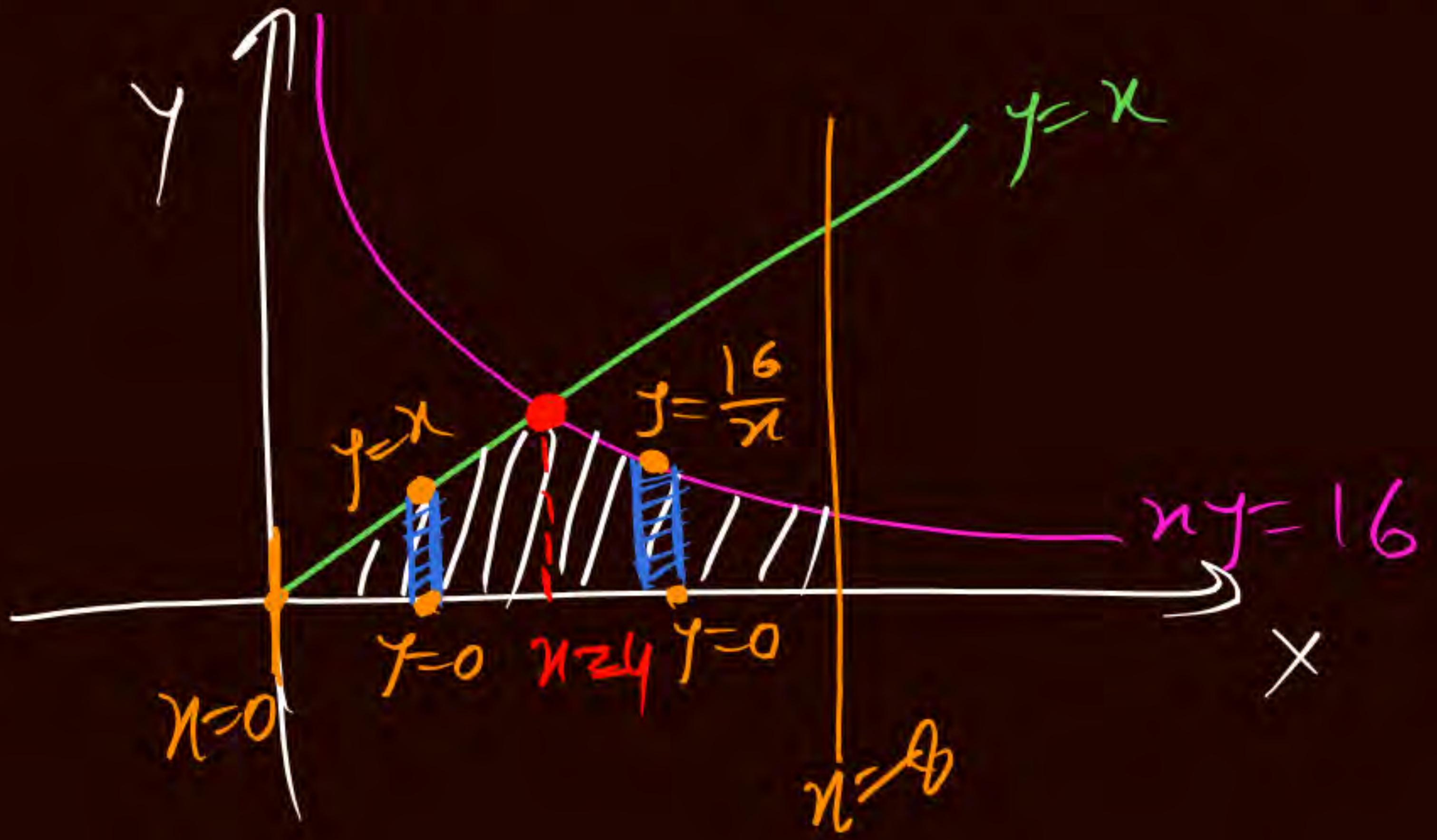
$$= \int_{y=0}^{1} \left[2(1-y)^2 \right] \cdot y\,dy = \cdots\cdots = \frac{1}{6}$$

(iii) Also find area of shaded region using H-Strip.

$$\boxed{\text{Area} = \iint (1)\,dx\,dy} = \int_0^1 \int_0^{2(1-y)} (1)\,dx\,dy = \cdots\cdots = 1$$

Q. Find $\iint x^2\,dy\,dx = ?$ over the Region bounded by $xy=16$, $y=x$, $x=0$, $x=8$

Ans $= 448$



$$I = \int_{x=0}^{4}\int_{y=0}^{x} x^2\,dy\,dx + \int_{x=4}^{8}\int_{y=0}^{16/x} x^2\,dy\,dx$$

$= \dots\dots = 448$

THANK - YOU