UNIT-IV

Regression & Functions of R.V

(a) Fit a St line $y = \bar{a}x + b$

(b) Fit a parabola $y = ax^2 + bx + c$

(c) Regression eqn of X on Y

(d) Regression eqn of Y on X

(e) Karl pearson correlation coefficient (r)

(F) Functions of Random variables

$$Y = T(X)$$

X R.V → pdf $f_X(x)$ → cdf $F_X(x)$

$Y = T(X)$ R.V → pdf $f_Y(y)$ → cdf $F_Y(y)$

Fit a St line:

$$y = ax + b \rightarrow (1)$$

Now its Normal eqns

$$(1)\ \Sigma y = a\Sigma x + nb$$

② $\Sigma xy = a\Sigma x^2 + b\Sigma x$

Solve above two eqns we get $a =$
$b =$

Sub.. a, b values in ①

we get required St line

<u>Fit a parabola</u>

$$y = ax^2 + bx + c \rightarrow ①$$

Now its Normal eqns

① $\Sigma y = a\Sigma x^2 + b\Sigma x + nc$

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

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

Solve above Three eqns we get $a =$
$b =$
$c =$

Sub.. a, b, c values in eqn ①

we get required parabola

① Fit a St line to The given data

[illegible]	13	16	17	20	25

x	10	12	13	16			
y	10	22	24	27	29	33	37

St line $y = ax + b$

Now its normal eqns

$$\left\{\begin{aligned} \sum y &= a\sum x + nb \\ \sum xy &= a\sum x^2 + b\sum x \end{aligned}\right\}$$

Here $n = 7$

x	y	xy	x^2
10	10	100	100
12	22	264	144
13	24	312	169
16	27	432	256
17	29	493	289
20	33	660	400
25	37	925	625
$\sum x$ 113	$\sum y$ 182	$\sum xy$ 3186	$\sum x^2$ 1938

values in above normal eqns

Subs these values in

$$182 = a(113) + b(7)$$

$$3186 = a(1938) + b(113)$$

Solving above two eqns we get

$$a = 1.56, \quad b = 0.82$$

Hence required St line

$$\boxed{y = 1.56x + 0.82}$$

② Fit a parabola to the given data

x	0	1	2	3	4
y	1	1.8	1.3	2.5	6.3

so Parabola:

$$y = ax^2 + bx + c$$

Now its Normal eqns

Here $\boxed{n = 5}$

$$\Sigma y = a\Sigma x^2 + b\Sigma x + cn$$

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

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

x	y	x^2	x^3	x^4	xy	x^2y
0	1	0	0	0	0	0
1	1.8	1	1	1	1.8	1.8
2	1.3	4	8	16	2.6	5.2
3	2.5	9	27	81	7.5	22.5
4	6.3	16	64	256	25.2	100.8
Σx 10	Σy 12.9	Σx^2 30	Σx^3 100	Σx^4 354	Σxy 37.1	Σx^2y 130.3

Sub. these values in above normal eqns

$$\left.\begin{aligned}
12.9 &= a(30) + b(10) + c(5) \\
37.1 &= a(100) + b(30) + c(10) \\
130.3 &= a(354) + b(100) + c(30)
\end{aligned}\right\}$$

Solve these three eqns we get

$$a = 0.55$$

$$b = -1.07$$

$c = 1.42$

Hence required parabola

$$\boxed{y = 0.55x^2 - 1.07x + 1.42}$$

Regression of X on Y

$$\boxed{X - \bar{X} = r\frac{\sigma_x}{\sigma_y}(Y - \bar{Y})}$$

$\bar{X}$ = Mean of X

$\bar{Y}$ = Mean of Y

Regression coeff of X on Y

$$r\frac{\sigma_x}{\sigma_y} = b_{XY} = \frac{\Sigma XY}{\Sigma Y^2} \rightarrow (1)$$

Regression eqn of Y on X

$$\boxed{Y - \bar{Y} = r\frac{\sigma_y}{\sigma_x}(X - \bar{X})}$$

$\bar{X}$ = Mean of X

$\bar{Y}$ = Mean of Y

Y on X

Regression coeff of

$$r\frac{\sigma_y}{\sigma_x} = b_{yx} = \frac{\Sigma XY}{\Sigma X^2} \rightarrow (2)$$

Now (1) × (2)

$$r^2 = b_{xy} \times b_{yx} = \frac{[\Sigma XY]^2}{\Sigma X^2 \Sigma Y^2}$$

$$\boxed{r = \sqrt{b_{xy} \cdot b_{yx}} = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}}$$

where r = Karl pearson correlation coeff

$$-1 \leq r \leq 1$$

$r = 1$ x & y are in same direction

$r = -1$ x & y are in opposite Direction

$r = 0$ x & y are independent

(1) Given the bi-variate data

x	1	5	3	2	1	1	7	3
y	6	1	0	0	1	2	1	5

... of X on Y

Then find (1) Regression eq

(2) Regression eqn of Y on x

(3) Karl pearson c.c (r)

$$① \quad X - \bar{X} = r \frac{\sigma x}{\sigma y} (Y - \bar{Y})$$

$$② \quad Y - \bar{Y} = r \frac{\sigma y}{\sigma x} (X - \bar{X})$$

$$③ \quad r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \; \Sigma Y^2}}$$

sol Here $\bar{x} = 2.8$

$\bar{y} = 2$

x	y	X $x-\bar{x}$ $x-2.8$	Y $y-\bar{y}$ $y-2$	XY	X^2	Y^2
1	6	-1.8	4	-7.2	3.24	16
5	1	2.2	-1	-2.2	4.84	1
3	0	0.2	-2	-0.4	0.04	4
2	0	-0.8	-2	1.6	0.64	4
1	1	-1.8	-1	1.8	3.24	1
				0	3.24	0

1	2	-1.8	0	[illegible]	[illegible]	[illegible]
7	1	4.2	-1	-4.2	17.64	1
3	5	0.2	3	0.6	0.04	9
				$\Sigma XY = -10$	$\Sigma X^2 = 32.92$	$\Sigma Y^2 = 36$

$\Sigma XY = -10$ $\qquad \bar{X} = 2.8$

$\Sigma X^2 = 32.92$ $\qquad \bar{Y} = 2$

$\Sigma Y^2 = 36$

Regression eqn of x on Y

$$X - \bar{X} = r \frac{\sigma x}{\sigma y} (Y - \bar{Y})$$

$$r \frac{\sigma x}{\sigma y} = \frac{\Sigma XY}{\Sigma Y^2}$$

$$= \frac{-10}{36}$$

$$= -0.27$$

$$r\frac{\sigma_x}{\sigma_y} = -0.27$$

Hence required Reg eqn of X on Y

$$X - 2.8 = -0.27\,(Y - 2)$$

$$X = 2.8 - 0.27Y + 0.54$$

$$\boxed{X = -0.27Y + 3.34}$$

Regression eqn of Y on X

$$Y - \bar{Y} = r\frac{\sigma_y}{\sigma_x}(X - \bar{x})$$

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2}$$

$$= \frac{-10}{32.92}$$

$$= -0.303$$

Hence

$$\ldots 303)(X - 2.8)$$

$Y - 2 = (-0.303)($

$Y = 2 - 0.303X + 0.8484$

$$\boxed{Y = -0.303X + 2.8484}$$

Now

karl pearson correlation coeff

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \, \Sigma Y^2}}$$

$$= \frac{-10}{\sqrt{32.92 \times 36}}$$

$$\boxed{r = -0.29} \qquad \boxed{-1 \leq r \leq 1}$$

Here r is $-ve$

$\therefore$ X & Y are in opposite direction

=

① Heights of fathers and sons are given in inches

Ht of father	65	66	67	67	68	69	71	73
Ht of son	67	68	64	68	72	70	69	70

① St line $y = ax + b$

② Parabola $y = ax^2 + bx + c$

③ Reg eqn of X on Y

④ Reg eqn of Y on X

⑤ r

⑥ $\tan\theta = \frac{m_1 - m_2}{1 + m_1 m_2}$

Functions of Random variable

X is a R.V

$f_X(x)$ is p.d.f $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

$(X, f(x))$ C.P.D

$F(x)$ is C.D.F

$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(x)\,dx$

Define a functions of R.V of

$$\boxed{Y = T(x)}$$

Here Y is another R.V

$f_Y(y)$ is pdf $\int_{-\infty}^{\infty} f_Y(y)\,dy = 1$

$F_Y(y)$ is C.D.F

$$F_Y(y) = P(Y \leq y) = \int_{y=-\infty}^{y} f_Y(y)\,dy$$

$$\boxed{\frac{d}{dy} F_Y(y) = f_Y(y)}$$

1. $Y = aX + b$
2. $Y = X^2$
3. $\max\{X, Y\} = T(Z)$
4. $\min\{X, Y\} = T(Z)$
5.

⑤ sum $\{X, Y\}$

① Consider the function $Y = aX + b$, where a and b are constants. Then find C.D.F of Y and P.D.F of Y

X R.V
$f_X(x)$
$F_X(x)$

Sol Given $Y = aX + b$

C.D.F of Y is

$$F_Y(y) = P(Y \le y)$$

$$= P(aX + b \le y)$$

$$= P\left(X \le \frac{y-b}{a}\right)$$

$$= F_X\left(\frac{y-b}{a}\right)$$

$$\boxed{F_Y(y) = F_X\left(\frac{y-b}{a}\right)}$$ ✓

P.d.f of Y

$$f_Y(y) = \frac{d}{dy}(y)$$

$$\frac{d}{dY} F_Y(y) = f_Y(y)$$

$$\frac{d}{dy} F_X\left(\frac{y-b}{a}\right) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

$$\boxed{f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)}$$ ✓

② Find C.D.F and pdf of $Y = 2X+3$ where X is R.V

sol Given $Y = 2X + 3$

$$F_Y(y) = P(Y \le y)$$

$$= P(2X+3 \le y)$$

$$= P\left(X \le \frac{y-3}{2}\right)$$

CDF of Y

$$\boxed{F_Y(y) = F_X\left(\frac{y-3}{2}\right)}$$ ✓

Now $\frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y-3}{2}\right)$

pdf of Y $\boxed{f_Y(y) = \frac{1}{2} f_X\left(\frac{y-3}{2}\right)}$

③ Find the C.D.F and P.D.F of $\boxed{\text{Power function}}$

$Y = X^2$

Sol Given $Y = X^2$

$$F_Y(y) = P(Y \leq y)$$

$$= P(X^2 \leq y)$$

$$= P(X \leq \pm\sqrt{y})$$

$$= P(|X| \leq \sqrt{y})$$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y})$$

$$\boxed{F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})}$$

Now $\frac{d}{dY} F_Y(y) = \frac{d}{dY}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right]$

$$= \frac{1}{2\sqrt{y}} f_X \sqrt{y} + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})$$

$$\boxed{f_Y(y) = \frac{f_X\sqrt{y} + f_X(-\sqrt{y})}{2\sqrt{y}}}$$

4. Maximum of two independent random variables

let $W = \max\{X, Y\}$

Now its C.D.F

$$F_W(w) = P(W \leq w)$$

$$= P(\max(X, Y) \leq w)$$

$$= P\{(X \leq w) \cap (Y \leq w)\}$$

$$= F_{XY}(w)$$

$$\boxed{F_W(w) = F_X(w)\, F_Y(w)}$$

$$f_W(w) = \frac{d}{dw} F_W(w)$$

$$= \frac{d}{dw}\left[F_X(w)\, F_Y(w)\right]$$

$$\boxed{f_W(w) = f_X(w)\, F_Y(w) + F_X(w)\, f_Y(w)}$$

problem

5 Assume that $W = \max\{X, Y\}$ where X & Y are independent r.v with respective p.d.f's

$$f_X(x) = \lambda e^{-\lambda x} \quad ; \quad x \geq 0$$

$$f_Y(y) = \mu e^{-\mu y} \quad , \quad y \geq 0$$

where $\lambda > 0$, $\mu > 0$ what is p.d.f of W

Sol Since

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)\, dx$$

$$= \int^{x} \lambda e^{-\lambda x}\, dx$$

$$= \lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^x$$

$$= -\left[e^{-\lambda x} - 1 \right]$$

$$\boxed{F_X(x) = 1 - e^{-\lambda x}}$$

||ly $F_Y(y) = P(Y \leq y) = \int_{-\infty}^{x} f_Y(y)\, dx$

$$= \int_0^y \mu e^{-\mu y}\, dy$$

$$\boxed{F_Y(y) = 1 - e^{-\mu y}}$$

Now p.d.f of w is given by

$$f_W(w) = f_X(w)\, F_Y(w) + F_X(w)\, f_Y(w)$$

$$= \lambda e^{-\lambda w}(1 - e^{-\mu w}) + \mu e^{-\mu w}(1 - e^{-\lambda w})$$

$$\boxed{f_W(w) = \lambda e^{-\lambda w} + \mu e^{-\mu w} - (\lambda + \mu) e^{-(\lambda+\mu) w}}$$

⑥ Minimum of Two Independent Random Variables

let $W = \min\{X, Y\}$

Now $F_W(w) = P(W \le w)$

$$= P(\min(X, Y) \le w)$$

$$= P((X \le w, X \le Y) \cup (Y \le w, Y \le X))$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= F_X(w) + F_Y(w) - F_{XY}(w, w)$$

$$\boxed{F_W(w) = F_X(w) + F_Y(w) - F_X(w)\, F_Y(w)}$$

since X & Y independent

Now $f_W(w) = \dfrac{d}{dw} F_W(w)$

$$= f_X(w) + f_Y(w) - [f_X(w)\, F_Y(w) + F_X(w)\, f_Y(w)]$$

$$\boxed{f_W(w) = f_X(w)[1 - F_Y(w)] + f_Y(w)[1 - F_X(w)]}$$

problem

7) Assume That $W = \min\{X, Y\}$ where X & Y are independent random variables with respective p.d.f's

$$f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0$$

$$f_Y(y) = \mu e^{-\mu y} \quad y \geq 0$$

where $\lambda > 0$, $\mu > 0$ what is The Pdf of W

Sol: please solve

same of $W = \max\{X, Y\}$

Sum of Two independent Random variables

let $S = X + Y$

$$F_S(s) = P(S \leq s)$$

$$= P(X + Y \leq s)$$

$$= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{s-y} f_{XY}(x, y)\, dx\, dy$$

X & Y

X & Y independent

$$= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{s-y} f_X(x)\, f_Y(y)\, dx\, dy$$

$$= \int_{y=-\infty}^{\infty} \left[\int_{x=-\infty}^{s-y} f_X(x)\, dx \right] f_Y(y)\, dy$$

$$\boxed{F_S(s) = \int_{y=-\infty}^{\infty} F_X(s-y)\, f_Y(y)\, dy}$$

Now $\frac{d}{ds} F_S(s) = f_S(s)$

$$\therefore f_S(s) = \int_{y=-\infty}^{\infty} \frac{d}{ds} F_X(s-y)\, f_Y(y)\, dy$$

$$f_S(s) = \int_{y=-\infty}^{\infty} f_X(s-y)\, f_Y(y)\, dy$$

By convolution theorem

$$\boxed{f_S(s) = f_X(s) * f_Y(s)}$$