PAGE NO :
DATE : / /

Unit- I

Curve Fitting by Least Squares

If x is an independent variable and $y=y(x)$ then for a known values of x, we can find corresponding values of y. But if $y=y(x)$ is not known, but we have a data set

$$\{(x_i, y_i) ; 1 \le i \le n\} \quad \text{--- } ①$$

We would like to fit a best fit curve to this data.

→ Fitting of a straight line.

Let the best fit st line to the data ① be $y = a + bx$ — ②
where a & b are to be determined.

Let Y_i be the values of y corresponding to $x = x_i$ as obtained from ②

$$\text{i.e. } Y_i = a + bx_i \quad \text{--- } ③$$

$$\text{Let } S = \sum_{i=1}^{n} (y_i - Y_i)^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \quad ④$$

Aim: To determine a and b such that S is least.

$$\text{i.e. } \frac{\partial S}{\partial a} = 0 \quad \& \quad \frac{\partial S}{\partial b} = 0$$

from ④

$$\frac{\partial S}{\partial a} = \sum_{i=1}^{n} 2(y_i - a - bx_i)(-1) = 0$$

$$\Rightarrow y_i - a - bx_i = 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i = \sum_{i=n} a + bx_i$$

$$\sum y_i = na + b\sum x_i \quad \text{— (5)}$$

diff (1) wrt b.

$$\cancel{\sum \frac{\partial y_i}{\partial b} = \frac{\partial na}{\partial b} + \frac{\partial (b\sum x_i)}{\partial b}}$$

$$\cancel{0 = 0 + \sum x_i}$$

$$\cancel{\sum x_i = 0}$$

$$\frac{\partial s}{\partial b} = 0 \Rightarrow \sum 2(y_i - a - bx_i)(-x_i) = 0$$

$$\sum x_i y_i = a\sum x_i + b\sum x_i^2 \quad \text{— (6)}$$

(5) & (6) are simultaneous Eqns for a & b and are called Normal Equations for line (2).

Solve (5) & (6) to get the values of a & b, substitute in (2), we get the least square line that best fits the given data.

→ Questions:-

1) Find the line of best fit for the points (1, 1), (2, 3) and (3, 2)

x	y	xy	x^2
1	1	1	1
2	3	6	4
3	2	6	9

$\sum x_i y_i = 13$

$\sum x_i^2 = 14$

$\sum x_i = 6$

$\sum y_i = 6$

Normal Eqns:-

$\sum y_i = na + b\sum x_i$

$6 = 3a + 6b$

$\sum x_i y_i = a\sum x_i + b\sum x_i^2$

$13 = 6a + 14b$

$a = 1, \quad b = 1/2$

⇒ The line of best fit is

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

2)

x	y	xy		x^2
0	9	0	~~81~~	0
1	8	8	~~64~~	1
2	24	48	~~576~~	4
3	28	84	~~784~~	9
4	26	104	~~676~~	16
5	20	100	~~400~~	25
15	115	344		55

$115 = 6a + 15b$

$344 = 15a + 55b$

$a = 11.09 \quad b = 3.22$

The line of best fit is

$$y = 11.09 + 3.22x$$

→ Fitting a straight line by the method of Least Squares.

Eg:- The result of measurement of electric resistance R of a wire at various temperatures θ is given below. Find a solⁿ $R = a + b\theta$ that best fits the data.

θ	R	$R\theta$	θ^2
19	76	1444	361
25	77	1925	625
30	79	2370	900
36	80	2880	1296
40	82	3280	1600
45	83	3735	2025
50	85	4250	2500
245	562	19884	9307

$$\Sigma R_i = na + b\Sigma\theta_i$$

$$562 = 7a + b(245)$$

$$\Sigma R\theta = a\Sigma\theta + b\Sigma\theta^2$$

$$19884 = a(245) + b(9307)$$

$$a = 70 \qquad b = 0.2923$$

$$R = 70 + 0.2923\,\theta$$

→ Fitting Geometric Squares.

To fit curves of the form $y = ab^x$, $y = ae^{bx}$, $y = ax^b$

To fit: $y = ab^x$

$$\log y = \log a + x \log b.$$

$$Y = A + Bx$$

→ Questions:-

1. Fit a curve of the form $y = ae^{bx}$

x	y
5	133
6	55
7	23
8	7
9	2
10	2

$Y = \log y$

$$\log_e y = \log_e a + bx \log_e e$$

$$Y = A + bx$$

Y	x	Yx	x^2
4.89	5	24.45	25
4.007	6	24.042	36
3.135	7	21.945	49
1.945	8	15.56	64
0.693	9	6.237	81
0.693	10	6.930	100
15.363	45	99.164	355

$$\Sigma Y = nA + b\Sigma x$$

$$15.363 = 6A + 45b$$

$$\Sigma Yx = \Sigma x A + b\Sigma x^2$$

$$99.164 = 45A + 355b$$

$$A = 9.4427$$

$$b = -0.9176$$

$$Y = A + bx$$

$$Y = 9.4427 - 0.9176\,x.$$

$$A = \log_e a \quad \therefore \quad a = e^A = 12615.733.$$

$$y = 12615.73\,e^{-0.9176x}$$

2. Fit a least squares curve of the form $y = ab^x$ to the data

$$y = ab^x$$
$$\log_e y = \log_e a + x \log_e b$$
$$Y = A + Bx$$

y	Y	x	Yx	x^2
32	3.4657	0	0	0
47	3.8501	1	3.8501	1
65	4.1743	2	8.3486	4
92	4.5217	3	13.5651	9
132	4.8828	4	19.5312	16
190	5.2470	5	26.235	25
275	5.6167	6	33.7002	36
	31.7583	21	105.2302	91

$$\Sigma Y = nA + B\Sigma x.$$
$$31.7583 = 7A + 21B$$

$$\Sigma Yx = A\Sigma x + B\Sigma x^2$$
$$105.2302 = 21A + 91B$$

$$A = 3.4702$$
$$B = 0.3555$$

$A = \log_e a$ $\quad a = e^A = e^{3.4702} = 32.1431$

$B = \log_e b$ $\quad b = e^B = e^{0.3555} = 1.4269$

$$\therefore y = 32.1431 \times 1.4269^x$$

3. Fitting a Least Squares Parabola:

Given data $\{(x_i, y_i) ; 1 \le i \le x\}$, if we want to fit a parabola of the form $y = a + bx + cx^2$, the normal eq^ns are:

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

→ Questions

1. Fit a parabola to the data

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	2.3	16	64	256	9.2	36.8
10	8.9	30	100	354	21.1	66.3

$$8.9 = 5a + 10b + 30c$$
$$21.1 = 10a + 30b + 100c$$
$$66.3 = 30a + 100c + 354c$$

$a = 0.3057$ $\quad y = 1.0711$

$b = 1.6585$ $\quad + 0.4157x$

$c = -0.3071$ $\quad - 0.0214x^2$

$\therefore$ ~~$y = 0.3057 + 1.6585b - 0.3071c$~~

2. Fit a least squares curve of the form $y = a + cx^2$

~~$\log y = \log a + \log x + 2\log x$~~

Normal eqns :- $\Sigma y = na + b\Sigma x + c\Sigma x^2$

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

x	y	x^2	x^3	x^4	xy	x^2y
1	3.8	1	1	1	3.8	3.8
2.5	15.0	6.25	15.625	39.0625	37.5	93.75
3.5	26.0	12.25	42.875	150.0625	91	318.5
4.0	33.0	16	64	256	132	528
	77.8	35.5		446.125		944.05

$77.8 = 4a + 35.5c$

$944.05 = 35.5a + 446.125c$

$a = 2.2789$

$b = 1.9347$

$y = 2.2789 + 1.9347x^2$

3. Fit a least squares geometric curve $y = ax^b$.

$\log y = \log a + b\log x$

$Y = A + bX$

Normal Eqns : $\Sigma Y = nA + b\Sigma X$

$\Sigma YX = A\Sigma X + b\Sigma X^2$

x	y	$Y = \log y$	$X = \log x$	YX	X^2
1	0.5	-0.6931	0	0	0
2	2	0.6931	0.6931	0.4804	0.4804
3	4.5	1.5041	1.0986	1.6524	1.2069
4	8	2.0794	1.3863	2.8827	1.9218
5	12.5	2.5257	1.6094	4.0648	2.5901
		6.1092	4.7874	9.0803	6.1992

$6.1092 = 5A + 4.7874b$

$9.0803 = 4.7874A + 6.1992(b)$

$A = -0.6932$

$b = 2$

$A = \log a$; $a = \log^{-1}(A) = 0.4999$

$y = ax^b \Rightarrow y = 0.4999x^2$

4. The voltage V across a capacitor at time t is given by ~~[illegible]~~. Fit a curve of the form $v = ae^{kt}$, what will be the voltage at $t = 9$ sec.

t	v	$V = \log v$	Vt	t^2
0	150	5.0106	0	0
2	63	4.1431	8.2862	4
4	28	3.3322	13.3288	16
6	12	2.4849	14.9094	36
8	5.6	1.7227	13.7816	64
20		16.6935	50.306	120

$V = ae^{kt}$.

$\log v = \log a + kt\log e$

$V = A + kt$

$y = a + bx$

Normal Eqns : $\Sigma V = nA + k\Sigma t$

$\Sigma Vt = A\Sigma t + k\Sigma t^2$

$16.6935 = 5A + 20k$

$50.306 = 20A + 120k$

$A = 4.9855$

$K = -0.4117$

$A = \log a \Rightarrow a = 146.2767$

$v = 146.2767 e^{-0.4117 t}$

At $t = 9$ sec, $v = 3.5973$

(20) → For the Eqn : $y = a + b/x$

put $\frac{1}{x} = X$; $y = a + bX$

Now, Normal Eqns ; $\Sigma y = na + b\Sigma X$

$\Sigma yX = a\Sigma X + b\Sigma X^2$

→ Multiple Regression :

If x and y are independent variables and we want to fit in a curve of the form $z = a + bx + cy$

Normal Eqns are : $\Sigma z = na + b\Sigma x + c\Sigma y$

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

$\Sigma yz = a\Sigma y + b\Sigma xy + c\Sigma y^2$

1) Fit a curve $z = a + bx + cy$ to the given data.
Estimate the value of z for $x = 10$ & y 3.5

x	y	z	x^2	y^2	xy	xz	yz
3	16	90	9	256	48	270	1440
5	10	72	25	100	50	360	720
6	7	54	36	49	42	324	378
8	4	42	64	16	32	336	168
12	3	30	144	9	36	360	90
14	2	12	196	4	28	168	24
48	42	300	474	434	236	1818	2820

$300 = 6a + 48b + 42c$

$1818 = 48a + 474b + 236c$

$2820 = 42a + 236b + 434c$

$a = 61.4$ $b = -3.6461$ $c = 2.5385$

best fit curve : $z = 61.4 - 3.6461x + 2.5385y$

given : $x = 10$, $y = 3.5$

$\therefore z = 33.8237$

2) Given the table, find the squares linear regression Eqn of z on x & y

Weight	z :	29	24	25	35	25.5	34.5
Height	x :	148	124	130	140	132	155
Age	y :	8	6	8	10	10	12

Estimate the weight of a girl who is 9 years old & 137 cm tall.

$z = a + bx + cy$

→ Same as prev. Q

PAGE NO :
DATE : / /

→ Correlation.

It indicates the strength and direction of the linear relationship b/w two random variables.

Let x and y be two random variables such that an increase in the value of one leads to an increase or decrease in the value of the other. Then, x & y are said to be correlated.

If an increase in x leads to increase in y, then x & y are said to be positively correlated.

If an increase in x leads to decrease in y, then x & y are said to be negatively correlated.

Let $\{(x_i, y_i),\ 1 \le i \le n\}$ be n data points.

then $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n} = \dfrac{1}{n}\Sigma x_i$

$\bar{y} = \frac{1}{n}\Sigma y_i$; $\sigma_x^2 = \frac{1}{n}\Sigma(x_i - \bar{x})^2$ → variance of x

$\sigma_y^2 = \frac{1}{n}\Sigma(y_i - \bar{y})^2$ → variance of y

σ_x = S.D of x ; σ_y = S.D of y

$$r = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}$$

r is called the co-efficient of Correlation b/w x and y

PAGE NO :
DATE : / /

If $r = \pm 1$, x & y are said to be perfectly correlated.

If $r = 0$, x & y are said to be not correlated.

If $x_i - \bar{x} = X_i$ & $y_i - \bar{y} = Y_i$

then $\sigma_x^2 = \frac{1}{n}\Sigma X_i^2$, $\sigma_y^2 = \frac{1}{n}\Sigma Y_i^2$

$$r = \frac{\Sigma X_i Y_i}{n\sigma_x\sigma_y} = \frac{\Sigma X_i Y_i}{\sqrt{\Sigma X_i^2\ \Sigma Y_i^2}}$$

Result : $-1 \le r \le 1$.

$\dfrac{(\Sigma a_i b_i)^2}{\Sigma a_i^2\ \Sigma b_i^2} \le 1$ (general)

$\Rightarrow \dfrac{(\Sigma X_i Y_i)^2}{\Sigma X_i^2\ \Sigma Y_i^2} \le 1$

$\left|\dfrac{(\Sigma X_i Y_i)}{\sqrt{\Sigma X_i^2\ \Sigma Y_i^2}}\right| \le \pm 1 \Rightarrow |r| \le \pm 1$

$\Rightarrow -1 \le r \le 1$.

→ Questions :

① Calculate r for

x	y	X $x-\bar{x}$	Y $y-\bar{y}$	X^2 $(x-\bar{x})^2$	Y^2 $(y-\bar{y})^2$	XY
10	18	-10	-3	100	9	30
14	12	-6	-9	36	81	54
18	24	-2	3	4	9	-6
22	6	2	-15	4	225	-30
26	30	6	9	36	81	54
30	36	10	15	100	225	150
				280	630	252

$\bar{x} = 20$ $\bar{y} = 21$

PAGE NO:
DATE: / /

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

$$r = \frac{252}{\sqrt{280 \times 630}}$$

$$r = \frac{252}{420}$$

$$r = 0.6$$

* (2) If r is the correlation coefficient b/w x and y and $z = ax+by$, then prove that $r = \frac{\sigma_z^2 - (a^2\sigma_x^2 + b^2\sigma_y^2)}{2ab\sigma_x\sigma_y}$

Proof: $z_i = ax_i + by_i$

$\Rightarrow \bar{z} = a\bar{x} + b\bar{y}$

$\therefore z_i - \bar{z} = a(x_i - \bar{x}) + b(y_i - \bar{y})$

Squaring :-

$$\Sigma(z_i - \bar{z})^2 = a^2\Sigma(x_i - \bar{x})^2 + b^2\Sigma(y_i - \bar{y})^2 + 2ab\Sigma(x_i - \bar{x})(y_i - \bar{y})$$

Summing

$\cancel{n(z_i - \bar{z})^2 = a^2 n(x_i - \bar{x})^2 + b^2 n(y_i - \bar{y})^2 + 2abn(x_i - \bar{x})(y_i - \bar{y})}$

$\sigma_x^2 = \frac{1}{n}\Sigma(x_i - \bar{x})^2 \Rightarrow \Sigma(x_i - \bar{x})^2 = n\sigma_x^2$

PAGE NO:
DATE: / /

$$n\sigma_z^2 = a^2 n\sigma_x^2 + b^2 n\sigma_y^2 + 2abnr\sigma_x\sigma_y$$

$$\sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2abr\sigma_x\sigma_y$$

$$\therefore r = \frac{\sigma_z^2 - a^2\sigma_x^2 - b^2\sigma_y^2}{2ab\sigma_x\sigma_y} \quad \text{--- (1)}$$

→ Cases :-

i) If $a=1$ & $b=1$

$\Rightarrow z = x+y$

$$\Rightarrow r = \frac{\sigma_{x+y}^2 - \sigma_x^2 - \sigma_y^2}{2\sigma_x\sigma_y} \quad \text{--- (2)}$$

ii) If $a=1$, $b=-1$

$\Rightarrow z = x-y$

$$\Rightarrow r = \frac{\sigma_{x-y}^2 - \sigma_x^2 - \sigma_y^2}{-2\sigma_x\sigma_y} = \frac{\sigma_x^2 + \sigma_y^2 - \sigma_{x-y}^2}{2\sigma_x\sigma_y} \quad \text{--- (3)}$$

(3) If S.D's of x & y are 2 & 3 resply, $r = 0.4$, find σ_{x+y} & σ_{x-y}.

$\sigma_x = 2 \quad \sigma_y = 3 \quad r = 0.4$

$$0.4 = \frac{\sigma_{x+y}^2 - 4 - 9}{2(2)(3)}$$

$$4.8 = \sigma_{x+y}^2 - 4 - 9$$

$\sigma_{x+y}^2 = 4.8 + 4 + 9$

$\sigma_{x+y}^2 = 17.8$

$\sigma_{x+y} = 4.2$

PAGE NO :
DATE : / /

$$0.4 = \frac{2^2 + 3^2 - \sigma^2_{x+y}}{2(2)(3)}$$

$$4.8 = 4 + 9 - \sigma^2_{x-y}$$

$$\sigma^2_{x-y} = 13 - 4.8$$

$$\sigma^2_{x-y} = 8.2$$

$$\sigma_{x-y} = 2.8635$$

*→ Regression.

Given some data $\{(x_i, y_i) ; 1 \le i \le n\}$. If we fit in a st. line to this data by considering x as independent variable and y as dependent on x, the line is called the regression line of y on x.

Eqⁿ of line of regression of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x}) = r\frac{\sigma_y}{\sigma_x}(x - \bar{x}) \quad (1)$$

$= \frac{\sum XY}{\sum X^2}$

$b_{yx} = r\frac{\sigma_y}{\sigma_x}$ → slope of line of regression.

$= \frac{\sum xy}{\sum y^2}$

→ x on y.

$$x - \bar{x} = b_{xy}(y - \bar{y}) = r\frac{\sigma_x}{\sigma_y}(y - \bar{y}) \quad (2)$$

$b_{xy} = r\frac{\sigma_x}{\sigma_y}$ → reciprocal of slope of line of regression

PAGE NO :
DATE : / /

from (1) & (2)

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

$$r = \pm\sqrt{b_{yx}\, b_{xy}}$$

b_{xy} & b_{yx} both have same sign. If both are $-ve$, r is $-ve$. If both are $+ve$, r is $+ve$.

→ Co-efficient of Correlation and Lines of Regression

1) Find the co-efficient of correlation and the lines of regression for the given data.

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	X^2	Y^2	XY
1	2	-2	-3	4	9	6
2	5	-1	0	1	0	0
3	3	0	-2	0	4	0
4	8	1	3	1	9	3
5	7	2	2	4	4	4
				10	26	13

$\bar{x} = 3$
$\bar{y} = 5$

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \cdot \sum Y^2}} = \frac{13}{\sqrt{260}} = \frac{13}{16.1245}$$

$$r = 0.8062$$

$$b_{yx} = \frac{\sum XY}{\sum X^2} = \frac{13}{10} = 1.3$$

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

$(y - 5) = 1.3(x - 3)$

$$b_{xy} = \frac{\Sigma XY}{\Sigma Y^2} = \frac{13}{26} = 0.5$$

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

$$(x-3) = 0.5\,(y-5)$$

② If the lines of regression are $2x+3y+1=0$ & $x+6y-4=0$ compute $\bar{x}$, $\bar{y}$ and r

$2x + 3y = -1 \quad \times -2 \rightarrow \quad -4x - 6y = 2$

$\quad x + 6y = 4$

$\quad -3x = 6$

$\quad x = -2 \rightarrow \bar{x}$

$2(-2) + 3y = -1$

$3y = 3$

$y = 1 \Rightarrow \bar{y}$

① $\rightarrow$ $(y-1) = b_{yx}(x+2)$

② $\rightarrow$ $(x+2) = b_{xy}(y-1)$.

$2x + 3y = -1 \qquad 2x = -3y - 1$

$\cancel{3y = -2x - 1}$ $\qquad x = -\frac{3}{2}y - 1 \qquad b_{xy} = -\frac{3}{2}$

$\cancel{y = -\frac{2}{3}x - \frac{1}{3}}$

$x + 6y = 4$

$6y = -x + 4 \qquad b_{yx} = -1/6$

$y = -\frac{1}{6}x + 4$

$$r^2 = b_{yx}\, b_{xy} = -\frac{1}{6} \times -\frac{3}{2} = \frac{1}{4}$$

$r = \frac{1}{2} = -0.5$ (b_{xy} & b_{yx} are $-ve$)

③ From the data giving rainfall (x) & discharge in a river (y), find the line of regression of y on x.

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	X^2	XY
1.53	33.5	−0.922	−8.32	0.85	7.6710
1.78	36.5	−0.672	−5.52	0.4515	3.7094
2.6	40.0	0.148	−1.82	0.0219	0.2693
2.95	45.8	0.498	3.98	0.248	1.9820
3.42	53.5	0.968	11.68	0.937	11.3062
				2.7055	24.9379

$\bar{x} = 2.452$

$\bar{y} = 41.86$

$$b_{yx} = \frac{\Sigma XY}{\Sigma X^2} = \frac{24.9379}{2.7055} = 9.2181$$

y on x : $(y - 41.86) = 9.2181\,(x - 2.452)$

④ If lines of regression are $3x + 2y = 26$ & $6x + y = 31$, find $\bar{x}, \bar{y}$ & r.

$\bar{x} = -4$

$\bar{y} = -7$.

$3x = 26 - 2y$

$x = \frac{26}{3} - \frac{2}{3}y$

$6x + y = 31$

$y = 31 - 6x$.

$\cancel{b_{xy} = -\frac{2}{3}}$ / $\cancel{b_{yx} = -6}$

$2y = 26 - 3x$

$y = 13 - \frac{3}{2}x$.

$b_{yx} = -\frac{3}{2}$, $b_{xy} = -\frac{1}{6}$.

$6x = 31 - y$

$x = 31 - \frac{1}{6}y$.

$$r^2 = -\frac{3}{2} \times -\frac{1}{6} = \frac{1}{4}$$

$$r = 1/2 = \boxed{-0.5}$$

PAGE NO :
DATE : / /

5) Eight Students obtained the foll^n marks in Maths in physics. Find the co-eff of correlation for the marks b/w 2 subjects. If a student got 18 marks in Maths, Estimate for marks in physics.

M	P	$m = M - \bar{M}$	$p = P - \bar{P}$	m^2	p^2	mp
26	20	4.125	-0.125	17.0156	0.0156	
20	22	-1.876	1.875	3.5193		
15	13	-6.875	-7.125	47.2656		
25	18	3.125	-2.125	9.7656		
30	27	8.125	6.875	66.0156		
22	25	0.125	4.875	0.0156		
10	12	-11.875	-8.125	141.0156		
27	24	5.125	3.875	26.266		

$\bar{M} = 21.875$

$\bar{P} = 20.125$

PAGE NO :
DATE : / /

→ Angle b/w two Regression lines.

l or of y on x

$$y + \bar{y} = b_{yx}(x - \bar{x}) \; ; \quad m_1 = b_{yx} = r\frac{\sigma_y}{\sigma_x}$$

l o.r of x on y

$$x - \bar{x} = b_{xy}(y - \bar{y}) \; ; \quad m_2 = 1/b_{xy} = \frac{1}{r}\frac{\sigma_y}{\sigma_x}$$

If θ is the angle b/w the 2 lines, then

$$\tan\theta = \left|\frac{m_1 - m_2}{1 + m_1 m_2}\right| = \left|\frac{r\frac{\sigma_y}{\sigma_x} - \frac{1}{r}\frac{\sigma_y}{\sigma_x}}{1 + \frac{\sigma_y^2}{\sigma_x^2}}\right|$$

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

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

$$\tan\theta = \frac{(1 - r^2)\,\sigma_y \sigma_x}{|r|\,(\sigma_x^2 + \sigma_y^2)}$$

i) If $r = \pm 1 \Rightarrow \tan\theta = 0 \Rightarrow \theta = 0$ or π.
l. of. r are coincident / parallel.

ii) If $r = 0 \Rightarrow \tan\theta = \infty \Rightarrow \theta = \pi/2$
The l. of. r are $\perp^r$.

→ Questions.

1. Obtain the lines of r & hence find the co-eff of corelation for the data.

x	y	$X = x-\bar{x}$	$Y = y-\bar{y}$	X^2	Y^2	XY
1	9	-3	-2	9	4	6
2	8	-2	-3	4	9	6
3	10	-1	-1	1	1	1
4	12	0	1	0	1	0
5	11	1	0	1	0	0
6	13	2	2	4	4	4
7	14	3	3	9	9	9
				28	28	26

$\bar{x} = 4$
$\bar{y} = 11$

$b_{yx} = \frac{\Sigma XY}{\Sigma X^2} = \frac{26}{28} = 0.9285$

$b_{xy} = \frac{\Sigma XY}{\Sigma Y^2} = \frac{26}{28} = 0.9285$

l. of r of y on x

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

$\Rightarrow (y-11) = 0.9285(x-4)$

l. of r of x on y

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

$\Rightarrow (x-4) = 0.9285(y-11)$

$r^2 = b_{xy} b_{yx} \Rightarrow \boxed{r = 0.9285}$

→ Random Variables:

A random variable is the one whose value is determined by the outcome of a random experiment.

(DRV) (count) Discrete Random variable: It is the one whose set of assumed values are countable.
Eg:- Rolling a dice, picking a card, tossing a coin

(CRV) (measure) Continuous Random Variable: It is the one whose set of assumed values is uncountable.
Eg:- Height, time

→ Discrete Probability Distribution:-

Suppose a DRV X is the outcome of an experiment

If $P(X = x_i) = P_i$; $i = 1, 2, 3 \ldots$
{or $p(x_i)$}

where, (i) $P_i \geq 0 \ \forall i$
(ii) $\Sigma P_i = 1$,

then $\langle x_i, P_i \rangle$ constitutes a discrete probability distribution.

Eg:- Toss a coin 2 times. Let X = no. of heads.
X = {HH, HT, TH, TT}

x_i	0	1	2
$p(x_i)$	1/4	1/2	1/4

$\Sigma p_i = 1$

→ Questions :-

1) The prob density funⁿ of a R.V. X is

X :	0	1	2	3	4	5	6
$P(x)$:	K	3K	5K	7K	9K	11K	13K

Find: a) $P(x<4)$, $P(X \geq 5)$, $P(3 \leq x \leq 6)$

b) What is the min value of K such that $P(x \leq 2) > 0.3$

→ $\Sigma P(X) = 1 \Rightarrow K + 3K + 5K + 7K + 9K + 11K + 13K = 1$

$49K = 1$

$K = 1/49$

$P(x<4) \Rightarrow K + 3K + 5K + 7K$

$= 16K$

$= 16/49$

$P(x \geq 5) = 11K + 13K = 24K = 24/49$

$P(3 \leq x \leq 6) = 9K + 11K + 13K = 33K = 33/49$

$P(x \leq 2) = K + 3K + 5K = 9K$

$9K > 0.3$

$K > \frac{0.3}{9}$

$K > \frac{1}{30}$

→ Mathematical Expectation :-

Consider a DRV X and a funⁿ $\phi(x)$ of X. Let $\{P(x_i)\}$ be a probability distributⁿ of X.

Then we define $E\{\phi(x)\} = \sum_i \phi(x_i) P(x_i)$ be the expected value of $\phi(x)$ — ①

* Mean of X : $E[X] = \sum_i x_i P(x_i)$ — ②

* Variance of X : $\sigma^2 = \sum_i (x_i - \mu)^2 P(x_i)$ — ③

$= \sum_i (x_i^2 + \mu^2 - 2\mu x_i) P(x_i)$

$\sigma^2 = \Sigma x_i^2 P(x_i) + \mu^2 \Sigma P(x_i) - 2\mu \Sigma x_i P(x_i)$

$\sigma^2 = \Sigma x_i^2 P(x_i) + \mu^2 - 2\mu(\mu)$

$\sigma^2 = \Sigma x_i^2 P(x_i) - \mu^2$

$\sigma^2 = E[x^2] - \{E(x)\}^2$ — ④

→ Questions.

1) Find the mean & Std. Deviatⁿ for the follⁿ.

x_i :	-5	-4	1	2
$P(x_i)$:	1/4	1/8	1/2	1/8

Mean $= \mu = E[X] = \sum_i x_i P(x_i)$

$\mu = -5(1/4) - 4(1/8) + 1(1/2) + 2(1/8)$

$\mu = -1$

Variance : $E[X^2] = \sum_i x_i^2 p(x_i)$

$= (-5)^2(1/4) + (-4)^2(1/8) + 1^2(1/2) + 2^2(1/8)$

$= 25/4 + 16/8 + 1/2 + 4/8$

$= 37/4$

$\sigma^2 = E[X^2] - \{E[X]\}^2$

$= 37/4 - (1)^2 = 33/4$

$\sigma = \sqrt{\frac{33}{4}} = 2.8723.$

2) Find μ & σ for.

x_i :	10	20	30	40
$p(x_i)$:	1/8	3/8	3/8	1/8

$\mu = E[X] = \sum_i x_i p(x_i)$

$= 10(1/8) + 20(3/8) + 30(3/8) + 40(1/8)$

$\boxed{\mu = 25}$

$E[X^2] = \sum_i x_i^2 p(x_i)$

$= 10^2(1/8) + 20^2(3/8) + 30^2(3/8) + 40^2(1/8)$

$= 700$

$\sigma^2 = E[X^2] - \{E[X]\}^2 = 700 - 25^2$

$= 75$

$\sigma = \sqrt{75} = 8.6602$

3). The prob distrib^n of a R.V. X is

x_i :	0	1	2	3	4	5	6	7
$p(x_i)$:	0	K	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

Find K, $P(X<6)$, $P(X\geq 6)$ & μ & σ

$\sum_i p(x_i) = 1.$

$0 + K + 2K + 2K + 3K + K^2 + 2K^2 + 7K^2 + K = 1.$

$10K^2 + 9K - 1 = 0$

$K = 0.1, -1.$

$\boxed{K = 0.1}$

$P(X<6) = 0 + K + 2K + 2K + 3K + K^2.$

$= K^2 + 8K$

$= 0.01 + 0.8$

$= 0.81$

$P(X\geq 6) = 2K^2 + 7K^2 + K.$

$= 0.02 + 0.07 + 0.1$

$= 0.19$

(or) $P(X\geq 6) = 1 - P(X<6)$

$= 1 - 0.81$

$= 0.19$

$\mu = E(X) = \sum_i x_i p(x_i)$

$= 1(K) + 2(2K) + 3(2K) + 4(3K) + 5(K^2)$

$+ 6(2K^2) + 7(7K^2+K)$

$= 30K + 66K^2$

$= 3 + 0.66$

$= 3.66$

PAGE NO :
DATE : / /

$E(x^2) = \sum_i x_i^2\, P(x_i)$

$= 1^2(k) + 2^2(2k) + 3^2(2k) + 4^2(3k) + 5^2(k^2) + 6^2(2k^2) + 7^2(7k^2 + k)$

$= 124\,k + 440\,k^2$

$= 12.4 + 4.4$

$= 16.8$

$\sigma^2 = E(x^2) - \{E(x)\}^2$

$= 16.8 - (3.66)^2$

$= 16.8 - 13.3956$

$= 3.4044$

$\sigma = \sqrt{3.4044}$

$\sigma = 1.8451$

→ Continuous R.V.

Given X is a random variable and $f(x)$ is the associated prob, then

i) $f(x) \geq 0$

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

iii) $P(a < x < b) = \int_a^b f(x)\,dx$

→ Questions :-

1) Is the f^n given by $f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$ a densitity for CRV X. Find $P(1 \leq X \leq 2)$

PAGE NO :
DATE : / /

i) $e^{-x} \geq 0 \;\forall\; x > 0$.

ii) $\int_{-\infty}^{\infty} f(x)\,dx = \cancel{\int_{-\infty}^{0} e^{-x}\,dx} + \int_0^{\infty} e^{-x}\,dx$

$= \left[\dfrac{e^{-x}}{-1}\right]_0^{\infty}$

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

$= (-1)\{0 - 1\}$

$= 1$

∴ It defines a density prob f^n for CRV.

iii) $\int_1^2 e^{-x}\,dx = \left.\dfrac{e^{-x}}{-1}\right|_1^2$

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

$= (-1)\{0.1353 - 0.3678\}$

$= 0.2325$

2) A random variable x has the density f^n $p(x) = \dfrac{k}{1+x^2}$, $-\infty < x < \infty$. Determine k and find i) $P(x \geq 0)$ ii) $(0 < x < 1)$.

$\int_{-\infty}^{\infty} \dfrac{k}{1+x^2}\,dx = 1$.

PAGE NO :
DATE : / /

$$k \tan^{-1}x \Big|_{-\infty}^{\infty}$$

$$k\{\tan^{-1}\infty - \tan^{-1}(-\infty)\} = 1$$

$$k\left\{\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right\} = 1$$

$$k(\pi) = 1$$

$$k = \frac{1}{\pi}$$

i) $P(x \geq 0)$

$$= \int_0^{\infty} \frac{k}{1+x^2}\,dx$$

$$= k \tan^{-1}x \Big|_0^{\infty}$$

$$= k\{\tan^{-1}\infty - \tan^{-1}0\}$$

$$= k\left\{\frac{\pi}{2} - 0\right\}$$

$$= \frac{1}{\pi}\left\{\frac{\pi}{2}\right\}$$

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

ii) $P(0 < x < 1)$

$$= \int_0^1 \frac{k}{1+x^2}\,dx.$$

PAGE NO :
DATE : / /

$$= k \tan^{-1}x \Big|_0^1$$

$$= k\{\tan^{-1}(1) - \tan^{-1}(0)\}$$

$$= k\{\pi/4 - 0\}$$

$$= \frac{1}{\pi}\left(\frac{\pi}{4}\right)$$

$$= 1/4$$

3) Find the constant α s.t the funn $f(x) = \begin{cases} \alpha x^2, & 0 < x < 3 \\ 0, & \text{otherwise} \end{cases}$ is a PDF

i) $P(1 \leq X < 2)$

ii) $P(x \geq 1)$

iii) $P(X > 2)$

$$\alpha \int_0^3 x^2\,dx$$

$$\frac{x^3}{3}\Big|_0^3$$

$$9 - 0$$

$$-9 > 0$$

$$\int_0^3 \alpha x^2\,dx.$$

Not required

$$= \alpha \frac{x^3}{3}\Big|_0^3$$

$$= \alpha(9 - 0)$$

$$= 9\alpha > 0.$$

$\therefore f(x)$ is greater than 0 $\forall$ $0 < x < 3$

$\therefore$ it is PDF.

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

$$\int_{-\infty}^{0} \cancel{f(x)\,dx} + \int_0^3 f(x)\,dx + \int_3^{\infty} \cancel{f(x)\,dx} = 1$$

$$\alpha \frac{x^3}{3}\Big|_0^3 = 1.$$

$$\alpha = 1/9$$

PAGE NO:
DATE: / /

i) $P(1 \le x < 2)$

$$= \alpha \int_1^2 x^2\,dx$$

$$= \alpha \frac{x^3}{3}\Big|_1^2$$

$$= \alpha\left(\frac{8}{3} - \frac{1}{3}\right)$$

$$= \frac{1}{9} \times \frac{7}{3} = 7/27$$

ii) $P(x \le 1)$

$$= \alpha \int_0^1 x^2\,dx$$

$$= \alpha \frac{x^3}{3}\Big|_0^1$$

$$= \alpha \left\{\frac{1}{3}\right\}$$

$$= \frac{1}{27}$$

iii) $P(x > 2) = \alpha \int_2^3 x^2\,dx$

$$= \alpha \frac{x^3}{3}\Big|_2^3$$

$$= \alpha \left\{\frac{27}{3} - \frac{8}{3}\right\}$$

$$= 19/27$$

PAGE NO:
DATE: / /

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

$$= \alpha \int_0^3 x^3\,dx$$

$$= \frac{x^4}{4}\Big|_0^3$$

$$= 81/4$$

4) If a R.V x has a pdf

$$f(x) = \begin{cases} ce^{-2x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

Find the prob that x takes a value b/w
i) 1 to 3
ii) less than 0.5

Finding c

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

$$\Rightarrow \int_{-\infty}^{0} 0\,dx + \int_0^{\infty} c\,e^{-2x}\,dx = 1$$

$$c \frac{e^{-2x}}{-2}\Big|_0^{\infty} = 1$$

$$\frac{c}{2}\left(e^{-\infty} - e^{-0}\right) = 1$$

$$\frac{-c}{2} = 1$$

$$c = -2$$

PAGE NO :
DATE : / /

i) $= \int_1^2 c\, e^{-2x}\, dx$

$= c \left. \frac{e^{-2x}}{-2} \right|_1^2$

$= \cancel{(+2)} \frac{(e^{-4} - e^{-1})}{\cancel{-2}}$

$= 0.0183 - 0.3678$

$= \cancel{+}0.3495$? (0.117)

ii) $\int_0^{0.5} c\, e^{-2x}\, dx$

$= \cancel{(+2)} \left. \frac{e^{-2x}}{\cancel{(-2)}} \right|_0^{0.5}$

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

$= 0.3678 - 1$

$= \cancel{+}0.632\cancel{2}$

5) The diameter of an electric cable is a CRV with $f(x) = kx(1-x)$; $0 \leq x \leq 1$ and otherwise, Find k, μ & σ.

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

$\int_0^1 kx(1-x)\, dx = 1$

PAGE NO :
DATE : / /

$k \int_0^1 x - x^2\, dx = 1$

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

$k \left\{ \frac{1}{2} - \frac{1}{3} \right\} = 1$

$\frac{k}{6} = 1$

$\boxed{k = 6}$

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

$\mu = k \int_0^1 x^2 (1-x)\, dx$

$\mu = 6 \int_0^1 x^2 - x^3\, dx$

$\mu = 6 \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^1$

$\mu = 6 \left\{ \frac{1}{3} - \frac{1}{4} \right\}$

$\mu = \frac{6}{12} = 0.5$

Variance $= \int_{-\infty}^{\infty} x^2 f(x)\, dx$

$= k \int_{0}^{1} x^3 - x^4\, dx$

PAGE NO :
DATE : / /

$$= 6\left\{\frac{x^4}{4} - \frac{x^5}{5}\right\}_0^1$$

$$= \frac{6}{20}$$

$\sigma^2 =$ Variance $- \mu^2$

$$\sigma^2 = \frac{6}{20} - (0.5)^2$$

$\sigma^2 = 0.30 - 0.25$

$\sigma^2 = 0.05$

$\boxed{\sigma = 0.2236}$

$\boxed{\mu = 0.5}$

$\boxed{K = 6}$

6. The pdf of X is given by $p(x) = y_0 e^{-|x|}$, $-\infty < x < \infty$. Find y_0, μ & variance.

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \qquad |x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$$

$$\int_{-\infty}^{\infty} y_0 e^{-|x|}\,dx = 1$$

$$\cancel{y_0 \frac{e^{-|x|}}{-1}\Big|_{-\infty}^{\infty}} \cancel{= 1}$$

$-x$ $\cancel{-y_0\{e\}}$

PAGE NO :
DATE : / /

$$= y_0 \int_{-\infty}^{0} e^{+x}\,dx + y_0 \int_0^{\infty} e^{-x}\,dx = 1$$

$$= y_0 \frac{e^{+x}}{+1}\Big|_{-\infty}^{0} + y_0 \frac{e^{-x}}{-1}\Big|_0^{\infty} = 1$$

$\cancel{= 1 - y_0(1)}$

$$= y_0\{1 - 0\} - y_0\{0 - 1\} = 1$$

$y_0 + y_0 = 1$

$2y_0 = 1$

$\boxed{y_0 = 1/2}$

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

$$= y_0 \int_{-\infty}^{0} x e^{x}\,dx + y_0 \int_0^{\infty} x e^{-x}\,dx$$

$$= y_0\left\{ x e^{x} - e^{x}\Big|_{-\infty}^{0} + \frac{x e^{-x}}{-1} - e^{-x}\Big|_0^{\infty} \right\}$$

$$= y_0\{0 - \cancel{1} - 0 + 0 - 0 - 0 + \cancel{1}\}$$

$= 0$

$$\text{Variance} : \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$= y_0\left\{ \int_{-\infty}^{0} x^2 e^{x}\,dx + \int_0^{\infty} x^2 e^{-x}\,dx \right\}$$

$x^2 e^{x} - 2\int e^{x} x\,dx$
$x^2 e^{x} - 2x e^{x} + 2e^{x}\Big|_{-\infty}^{0}$

$-x^2 e^{-x} + 2\int e^{-x} x$
$-x^2 e^{-x} - 2x e^{-x} + 2e^{-x}$

PAGE NO :
DATE : / /

$$= y_0 \left\{ x^2 e^{x} - 2xe^{x} + 2e^{x} \Big|_{-\infty}^{0} + \left(-x^2 e^{-x} + 2xe^{-x} - 2e^{-x} \right) \Big|_{0} \right.$$

$$= y_0 \left\{ 2 + 2 \right\}$$

$$= y_0 (4) = \frac{1}{\cancel{2}} \cancel{(4)}^{2}$$

$$= 2$$

$$\sigma^2 = E(x^2) - E(x)^2$$

$$= 2 - 0$$

$$= 2$$

$$\sigma = \sqrt{2}$$

→ Probability Distributions

A p.d gives the prob of an event in an experiment.

1) Discrete p.d

2) continuous p.d

- Binomial Distribution

It is a discrete p.d that given the prob of obtaining exactly x successess in n trials of an experiment.

$$B(n, p, x) = {}^nC_x \, p^x q^{n-x}$$

p: prob of successes in a single trial of the event.

$${}^nC_x = \frac{n!}{x!(n-x)!} \qquad q = 1-p$$

PAGE NO :
DATE : / /

→ Questions.

1) In an unbiased coin out 10 tosses, what is the prob of getting i) 2 heads ii) 6 heads iii) less than 3 heads, iv) At least 3 heads.

→ $p = q = \frac{1}{2}$, $n = 10$

i) $x = 2$

$$B\left(10, \tfrac{1}{2}, 2\right) = {}^{10}C_2 \left(\tfrac{1}{2}\right)^2 \left(\tfrac{1}{2}\right)^8$$

$$= \frac{10!}{2! \times 8!} \left(\frac{1}{4}\right)\left(\frac{1}{256}\right) \quad \left(\frac{10 \times 9 \times \cancel{8!}}{2! \times \cancel{8!}}\right)$$

$$= \frac{45}{4} \times \frac{1}{256}$$

$$= 0.04394$$

ii) $x = 6$.

$$B\left(10, \tfrac{1}{2}, 6\right) = {}^{10}C_6 \left(\tfrac{1}{2}\right)^6 \left(\tfrac{1}{2}\right)^4$$

$$= \frac{10 \times 9 \times 8 \times 7 \times \cancel{6!}}{\cancel{6!} \times 4!} \times \frac{1}{1024}$$

$$= 0.2050$$

iii) $x < 3 \Rightarrow P(x=0) + P(x=1) + P(x=2)$

PAGE NO :
DATE : / /

$$B(10, \tfrac{1}{2}, 0) = {}^{10}C_0 (p)^0 (q)^{10}$$

$$= \frac{10!}{0! \times 10!} (1) \left(\frac{1}{2}\right)^{10}$$

$$= \frac{1}{1024} = 0.0009$$

$$B(10, \tfrac{1}{2}, 1) = {}^{10}C_1 (p)^1 (q)^9$$

$$= \frac{10!}{1! \times 9!} \left(\frac{1}{2}\right)^{10}$$

$$= \frac{9}{1024} = 0.0087$$

$$B(10, \tfrac{1}{2}, 2) = {}^{10}C_2 \, p^2 q^8$$

$$= \frac{10!}{2! \times 8!} \left(\frac{1}{2}\right)^{10}$$

$$= \frac{\cancel{10}^{5} \times 9}{\cancel{2}} \times \frac{1}{1024}$$

$$= 0.0439$$

$$\therefore P(x < 3) \Rightarrow 0.0009 + 0.0087 + 0.0439$$

$$= 0.0535$$

iv) $x \geq 3$

$$= 1 - P(x<3)$$

$$= 1 - 0.0535$$

$$= 0.9465$$

PAGE NO :
DATE : / /

2) In a biased coin, where head appears thrice as often as tails, find the prob of getting 6 heads in 12 tosses.

$$p = \frac{3}{4} \qquad q = \frac{1}{4}$$

$$B(12, \tfrac{3}{4}, 6) = {}^{12}C_6 \left(\frac{3}{4}\right)^6 \left(\frac{1}{4}\right)^6$$

$$= \frac{12!}{6!\ 6!} \cdot \frac{729}{16777216}$$

$$= \frac{12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6}{6!} \times$$

$$= 0.0396 \times \frac{729}{6!}$$

$$= 0.04009$$

3. In a large consignment of items, 5% are defective. If random sample of 8 items is taken for inspection, what is the prob that it has one or more defective.

$$p = \frac{5}{100} = \frac{1}{20} \qquad q = \frac{19}{20} \qquad n = 8$$

$$P(x \geq 1) = 1 - P(x=0)$$

$$B(8, \tfrac{1}{20}, 0) = {}^{8}C_0 \left(\frac{1}{20}\right)^0 \left(\frac{19}{20}\right)^8$$

$$= 1 \times 1 \times 0.6634$$

$$= 0.6634$$

$\frac{20 \times 19!}{1 \times 19!}$

$P(x \geq 1) = 1 - 0.6634$

$= 0.3366$

4) In sampling a large no. of machine parts, the mean no. of defectives in a sample of 20 is 2. Out of 1000 such such samples, how many would be expected to contain at least 3 defective parts.

$p = \frac{2}{20} = \frac{1}{10}$ $q = \frac{9}{10}$

for $n = 20$

$B(20, 1/10, 3 \text{ or more})$

$= 1 - [P(x=0) + P(x=1) + P(x=2)]$

$= 1 - \left[{}^{20}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{20} + {}^{20}C_1 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^{19} + {}^{20}C_2 \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^{18}\right]$

$= 1 - [0.3486 + 20 \times 0.1 \times 0.3874 + 190 \times 0.01 \times 0.4304]$

$= 1 - [0.3486 + 0.7748 + 0.8187]$

$p = 0.01$

$n = 1$, $r = 0.5$

5) If the prob of success in a single tried is 0.01 how many trials are necessary in order that the prob. of at least one success is 0.5 or more.

6) In a Binomial distribution, the mean is 6 and variance is $\frac{2}{?}$. Find the
i) total no. of trials.
ii) Prob. of 4 success.

$\mu = np = 6$ $\sigma^2 = npq = 2$

$\frac{\cancel{np}q}{\cancel{np}} = \frac{2}{6} \Rightarrow q = \frac{1}{3}$

$\therefore p = 2/3$

$np = 6$

$n(2/3) = 6^{3}$

i) $\boxed{n = 9}$

ii) $x = 4$.

$B(9, 2/3, 4) = {}^9C_4 \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^5$

$= \frac{9 \times 8 \times 7 \times 6 \times 5 \times \cancel{4!}}{\cancel{4!}\ 5!} \times \frac{16}{19683}$

$= \frac{24192}{236196} = 0.1024$

PAGE NO:
DATE: / /

7) In a bombing action, there is a 50% chance that a bomb will strike the target. Two direct hits are needed to destroy the target completely. How many bombs are required to be dropped to give a 99% or better chance of completely destroy the target.

→ $p = q = \frac{1}{2}$, let the no. of bombs dropped be n.

$P(x=0) = {}^nC_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^n$

$P(x=1) = {}^nC_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{n-1}$

given :- $1 - \left[{}^nC_0 \left(\frac{1}{2}\right)^n + {}^nC_1 \left(\frac{1}{2}\right)^n \right] \geq 99$

$(1+n)\left(\frac{1}{2}\right)^n \not< 0.01$

$n = 11$

→ Poison Distribution.

$\mu = np$ $P(\mu, x) = \frac{e^{-\mu}\mu^x}{x!}$

mean & variance of P.D = μ.

→ Questions

1) A Poisson variate is such that
$P(X=2) = 9P(X=4) + 90P(X=6)$
Find the mean & variance of the distribution.

PAGE NO:
DATE: / /

$\frac{e^{-\mu}\mu^2}{2!} = \frac{9e^{-\mu}\mu^4}{4!} + \frac{90e^{-\mu}\mu^6}{6!}$

$\frac{1}{2!} = \mu^2\left(\frac{9}{4\times3\times2\times1} + \frac{90\mu^2}{6\times5\times4\times3\times2\times1}\right)$

$\mu^4 + 3\mu^2 - 4 = 0$

$\mu^2 = 1, -4$ $\mu^2 = 1$

2) Suppose 2% of the items produced by a machine are defective. Find the prob that there are 5 defective items in a sample of 200 items.

$n = 200$ $x = 5$ $p = \frac{2}{100} = \frac{1}{50}$ $q = \frac{49}{50}$

$\mu = np = 200 \times \frac{1}{50} = 4$

$P(4, 5) = \frac{e^{-4}(4)^5}{5!}$

$= 0.1563$

3) A car hire firm has 2 cars which it rents out. The demand for a car on each day follows poisson dist. with mean 1.5. Calculate the proportion of days on which there is
i) no demand
ii) the demand is refused.

PAGE NO :
DATE : / /

→ $\mu = 1.5$

i) $x = 0$

$P(1.5, 0) = \frac{e^{-1.5} (1.5)^0}{0!}$ $n=2$

$= 0.2231$

ii) $1 - \{P(x=0) + P(x=1) + P(x=2)\}$

$1 - \left\{0.2231 + \frac{e^{-1.5} \times 1.5}{1} + \frac{e^{-1.5} \times (1.5)^2}{2}\right\}$

$1 - \{0.2231 + 0.3346 + 0.2509\}$

$= 1 - 0.8086$

$= 0.1914$

4) If the prob of an item being produced defective by a machine is 0.002 and its are packed in sets of 10. the p.d. to calculate the approximate no. of packets containing no defective and one defective item in a consignment of 100000 packets.

→ $n = 10$, $p = 0.002$ 10×0.002

$\mu = np = 0.02$

PAGE NO :
DATE : / /

$P(x=0) = \frac{e^{-0.02} (0.02)^0}{0!}$

$= ~~0.98~~ \; 0.9802$

$P(x=1) = \frac{e^{-0.02} (0.02)^1}{1!}$

$= ~~0.01637~~ \; 0.0196$

~~$P(x=0) + P(x=1) = 0.9824$~~

~~Prob of a set of 10~~

Out of 10000 packets, ~9802 will have no defective item.

" " " " , ~196 will have 1 defective item.

5) If the prob of a bad reaction from a certain medicine is 0.001, determine the chance that out of 2000 ppl, more than 2 will get a bad reaction.

$p = 0.001$ $n = 2000$

$\mu = np = 2$

$P(x>2) = 1 - \{P(x=0) + P(x=1) + P(x=2)\}$

$= 1 - \left\{\frac{e^{-2}(2)^0}{0!} + \frac{e^{-2}(2)^1}{1!} + \frac{e^{-2}(2)^2}{2!}\right\}$

$= 1 - 0.6766$

$= 0.3233$