UNIT - II

0/10/23

STATISTICS AND REGRESSION ANALYSIS

• Simple linear Regression and corelation.

Linear Regression: The statistical method which helps to estimate the unknown value of one variable from the known value of related variable. It is called linear regression.

• The line discribed in the average relationship between two variables is called line regression.

i.e, $y = ax + b$

Methods to calculate Regression.

- Graphical Method
- Mathematical method

Methods of least squares:

We have 4 methods:

- Fitting a straight line. $y = ax + b$ (or) $y = a + bx$
- Fitting a parabola of 2nd degree.

  $y = a + bx + cx^2$ (or) $y = ax^2 + bx + c$
- Exponencial Curve $y = ae^{bx} \Rightarrow Y = A + BX$
- Power Curve $y = ab^x$

Fitting a straight line:

$y = ax + b$ is a straight line where, "a, b" are constants "a" indicates volume of "y" when "x" is zero (0).

a) <u>Regression Equation of "y" on "x"</u>: ($y = ax + b$ (or) $y = a + bx$

Let $y = ax + b$

apply Σ on both sides

$$\Sigma y = \Sigma (ax + b)$$

$$\Sigma y = a \Sigma x + b \Sigma (1) \quad — (1)$$

$\because \Sigma (1) = n$

$$\boxed{\Sigma y = a \Sigma x + nb}$$

(1) multiply by x on both sides

$$\boxed{\Sigma xy = a \Sigma x^2 + b \Sigma x}$$

These are normal equations.

b) <u>Regression Equation of "x" on "y"</u>:

Let $x = a + by$

apply Σ on both sides

$$\Sigma x = \Sigma (a + by)$$

$$\boxed{\Sigma x = na + b \Sigma y}$$

Multiply by y on both sides

$$\boxed{\Sigma xy = a \Sigma y + b \Sigma y^2}$$

These are normal equations.

Find Regression line of "y" on "x" and "x" on "y".
Determining the equation straight line which fits the data

x	10	12	13	16	17	20	25	/ 113	n = 7
y	10	22	24	27	29	33	37	/ 182	

1983
5188

a) Regression Eq. of 'y' on 'x':
The straight line to be fitted

$$y = a + bx$$

0.798
1.561

Take sigma on both sids

$$\Sigma y = \Sigma (a + bx)$$

$$\Sigma y = a \Sigma (1) + b \Sigma x \quad -①$$

$$\Sigma (1) = n$$

$$\boxed{\Sigma y = an + b \Sigma x} \quad -②$$

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

① multiply by x

$$\boxed{\Sigma xy = a \Sigma x + b \Sigma x^2} \quad -③$$

$3186 = a(113) + b(1983)$

Here, $\Sigma x = 113$ $\quad \Sigma x^2 = 1983$ $\quad \Sigma xy = 3186$

$\Sigma y = 182$ $\quad \Sigma y^2 = 5188$

② => $182 = 7a + 113b$

③ => $3186 = 113a + 1983b$

$a = 0.798$
$b = 1.561$

$\therefore a = 0.798, \; b = 1.561$

Required straight line $y = a + bx$

$$y = 0.79 + (1.56)x.$$

12/10/2023

Regression Equation of 'x' on 'y':

$$x = ay + b$$

$$\Sigma x = \Sigma (ay) + \Sigma b$$

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

$$\Sigma x = a \Sigma y + nb \quad — (4)$$

$$\Sigma xy = a \Sigma y^2 + b \Sigma y \quad — (5)$$

(4) => $113 = a(182) + 7b$ 0.54 2.

(5) => $3186 = a(5188) + 182(b)$

$a = 0.54$

$b = 2.$

Required straight line: ~~y = a~~ $x = ay + b$

$$x = (0.54)y + 2.$$

(H.W) Fit a straight line

x	11	7	2	5	8	6	10
y	7	5	3	2	6	4	8

$\Sigma x = 49$ $\Sigma x^2 = 399$

$\Sigma y = 35$ $\Sigma y^2 = 203$

$\Sigma xy = 280$ $4x = 5y + 3$

$a = 1.25 \left(\frac{5}{4}\right)$ $b = 0.75 \left(3/4\right)$ $x = \frac{5y}{4} + \frac{3}{4}$

Fitting A Parabola:

The equation $y = ax^2 + bx + c$. where, a, b, c are constants and x, y be variables.

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

Normal Equations: $\Sigma y = a \Sigma x^2 + b \Sigma x + nc$

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

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

Fit a 2nd degree parabola $y = ax^2 + bx + c$ in the least square method for the given data. $(y = a + bx + cx^2)$
Hence Estimate 'y' at $x = 6$.

x	1	2	3	4	5	$\Sigma x = 15$
y	10	12	13	16	19	$\Sigma y = 70$

$\Sigma x^2 = 55$ $\quad \Sigma x^3 = 225$ $\quad \Sigma x^4 = 979$

$1030 \; \Sigma y^2$

$\Sigma xy = 232$

$\Sigma x^2 y = 906.$

The parabola to be fitted:

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

Normal Equations:

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

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

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

① $\Rightarrow$ $70 = a(55) + b(15) + 5(c)$

② $\Rightarrow$ $232 = a(225) + b(55) + c(15)$

③ $\Rightarrow$ $906 = a(979) + b(225) + c(55)$

Solving, $a = 0.2857$

$b = 0.4857$

$c = 9.4$

$\therefore$ Required Parabola: $y = ax^2 + bx + c$

$$y = (0.28)x^2 + (0.48)x + 9.4$$

If $x = 6$, $\quad y = \cancel{41.028}\ 22.59.$

$\therefore$ at $x = 6$, $y = 22.599$

H.W Parabola

x	0	1	2	3	4
y	1	3	4	5	6

if $x = 1$.

13/10/23

(H.W) The parabola to be fitted:

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

Normal Equations: $\Sigma y = a\Sigma x^2 + b\Sigma x + nc$ — ①

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

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

Here, $\Sigma x = 10$ $\Sigma x^2 = 30$ $\Sigma x^3 = 100$ $\Sigma x^4 = 354$

$\Sigma y = 19$ $\Sigma x^2 y = 160$.

$\Sigma xy = 50$ $n = 5$.

① => $19 = a(30) + b(10) + c(5)$

② => $50 = a(100) + b(30) + c(10)$

③ => $160 = a(354) + b(100) + c(30)$

Solving, $a = -1/7 = -0.1428$

$b = 63/35 = 1.771$

$c = 39/35 = 1.1142$

∴ Required Parabola: $y = ax^2 + bx + c$

$$y = \left(\frac{-1}{7}\right)x^2 + \frac{63}{35}(x) + \frac{39}{35}$$

for $x = 1$

$$y = \frac{-1}{7}(1)^2 + \frac{63}{35}(1) + \frac{39}{35} = 2.77$$

∴ If $x = 1$, y will be $2.77. \cong 3$.

Exponencial Curve:

The curve of the form

$$y = a e^{bx}$$

Applying log on both sides

$$\log y = \log (a e^{bx})$$

$$\log y = \log a + \log e^{bx}$$

$$\log y = \log a + \cancel{\log}\ bx$$

Let $Y = \log y$, $A = \log a$

$$\Rightarrow Y = A + bx$$

Normal Equations:

Applying Σ on both sides,

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

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

Fit the curve of the form $y = a e^{bx}$

x	77	100	185	239	285	Σx 886	Σx^2 188500
y	2.4	3.4	7	11.1	19.6	ΣY 9.4	ΣxY 1968.09
Y	0.875	1.223	1.94	2.4	2.97		$n = 5$.

The curve of the form : $y = a e^{bx}$

$$\Rightarrow \log_e y = \log_e a + \log_e e^{bx}$$

$$\log y = \log a + bx$$

0.181
9.588

$\cancel{\Sigma y}$ $\cdot \Sigma Y = nA + b\Sigma x$ — (1)

$\Sigma xY = A\Sigma x + b\Sigma x^2$ — (2)

$\log_x = 2$

$a = x^2$

(1) $\Rightarrow$ $9.4 = 5(A) + b(886)$

(2) $\Rightarrow$ $1968.09 = 886(A) + b(188500)$

$\because A = \log a$

Solving, $A = 0.183 \Rightarrow a = 1.197$

$b = 9.602$

$a = A = e^{0.183}$

$\therefore$ Required Exponencial Curve:

$y = ae^{bx}$

$y = (1.197)e^{(9.6)x}$

H.W

Fit the curve (exponencial):

i)

x	0	1	2	3	4	5	6	7
y	10	21	35	59	92	200	400	610

$y = (10.49)e^{(0.58)x}$

ii)

x	0	1	2	3	4	5	6	7
y	20	30	52	77	135	211	326	550

$y = (19.46)e^{(0.474)x}$

17/10/2023

POWER CURVE: ($\log_{10}$)

The curve of the form $y = ab^x$

$$y = ab^x$$

$$\log_{10} y = \log_{10}(ab^x)$$

$$\log_{10} y = \log_{10} a + x \log b$$

Let $Y = \log_{10} y$; $A = \log_{10} a$; $B = \log_{10} b$

$$\Rightarrow Y = A + Bx.$$

Normal Equations: $\Sigma Y = nA + B\Sigma x$ —— ①

$\Sigma xY = A\Sigma x + B\Sigma x^2$ —— ②

Q Applying a relation of form $y = ab^x$ for the following data By the method of ~~to~~ least squares.

x	2	3	4	5	6
y	8.3	15.4	33.1	65.2	127.4

The power curve to be fitted:

The curve of form: $y = ab^x$

$$\log y = \log(ab^x)$$

$$\log y = \log(a b^x)$$

$$\log y = \log a + \log(b^x)$$

$$\log y = \log a + x \log b \quad - (1)$$

Let $\left.\begin{array}{l}\log y = Y \\ \log a = A \\ \log b = B\end{array}\right\} - (2)$

$$Y = A + xB$$

Normal Equations: $\sum Y = nA + B\sum x \quad - (3)$

$$\sum xY = A\sum x + B\sum x^2 \quad - (4)$$

x	2	3	4	5	6	20
y	8.3	15.4	33.1	65.2	127.4	
$Y = \log y$	0.919	1.187	1.519	1.814	2.1051	7.545 ~~6.972~~
xY	1.838	3.5625	6.07	9.07	12.6	~~30.25~~ 33.1819
x^2	4	9	16	25	36	90

Here, $\sum x = 20$ $\qquad \sum x^2 = 90$

$\sum Y = 7.545$

$\sum xY = 33.182$

(3) $\Rightarrow$ $7.545 = 5(A) + B(20)$

(4) $\Rightarrow$ $33.1819 = A(20) + B(90)$

Solving: $A = 0.3096$ $\qquad$ 0.3096
$\qquad$ 0.299

$B = 0.3$

(2) $\Rightarrow$ $A = \log_{10} a$

$0.3096 = \log_{10} a$

$a = 2.04$

$B = \log_{10} b$

$0.3 = \log_{10} b$

$b = 1.99$

$\therefore$ Required power curve :

$y = ab^x$

$y = (2.04)(1.99)^x$

Fit a power curve by the given data.

x	1	2	3	4	5	6
y	151	100	61	50	20	8

The power curve to be fitted :

The curve of form: $y = ab^x$

$\log y = \log (ab^x)$

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

Let, $\log y = Y$

$\log a = A, \log b = B.$ $\Big\}$ — (1)

$$Y = A + xB$$

Normal Equations: $\Sigma Y = nA + B\Sigma x$ — (2)

$$\Sigma xY = A\Sigma x + B\Sigma x^2 \quad -(3)$$

x	1	2	3	4	5	6
y	151	100	61	50	20	8
$Y\ (\log_{10} y)$	2.178	2	1.785	1.698	1.301	0.903
xY						
x^2						

Hence, $\Sigma x = 21$ $\quad \Sigma xY = 30.2545$

(2) $\Sigma Y = 9.86$ $\quad \Sigma x^2 = 91$

(2) $\Rightarrow$ $9.86 = 6(A) + B(21)$

(3) $\Rightarrow$ $30.25 = A(21) + B(91)$

Solving, $A = 2.5$

$B = 0.24465$

(1) $\Rightarrow$ $A = \log_{10} a \Rightarrow a = 316.838$

$B = \log_{10} b \Rightarrow b = 0.569$

Required power curve: $y = ab^x$

$$y = (316.838)(0.569)^x$$

H.W Fit a power curve

x	0	1	2	3	4	5	6
y	10	21	35	~~59~~ 59	92	200	600

Properties of coefficient of correlation:

i) $-1 \leq r \leq 1$

ii) If $r = 1$. It is called positive correlation

i.e, the direction of variability is in the same direction. $(x\uparrow \Rightarrow y\uparrow)$ $(x\downarrow \Rightarrow y\downarrow)$

iii) If $r = -1$. It is called negative correlation.

i.e, the direction of variability is in opposite direction $(x\uparrow \Rightarrow y\downarrow)$ $(x\downarrow \Rightarrow y\uparrow)$

iv) If $r = 0$. It is un-correlated.

i.e, there is no relation.

Find If there is any significant correlation between the heights and weights given.

Heights (inches)	57	51	62	63	64	65	55	58	57
Weight	113	117	128	126	130	129	111	116	112

19/10/23

Karl Pearson's Coefficient of correlation

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$$

CORRELATION

If the change increasing or decreasing in one variable effects the change increasing or decreasing in another variable.

Then, ~~that~~ these two variables are said to be corelated.

Example: • Heights and Weights of group of people.

• Income and Expenses of a employee.

Types of corelation coefficients

• Karl Pearson's (r)

• Spearman's coefficient of correlation. (Rank correlation).

Karl Pearson's coefficient of correlation

This coefficient of correlation studies the variability between two linear variables.

It is denoted by "r".

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

$\bar{x} = \frac{\Sigma x_i}{n}$ (mean of x values)

$\bar{y} = \frac{\Sigma y_i}{n}$ (mean of y values)

$X = x - \bar{x}$

$Y = y - \bar{y}$

Karl Pearson's coefficient of correlation:

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

x	y	$X = x - \bar{x}$ (59)	X^2	$Y = y - \bar{y}$	Y^2	XY
55	58	-4	16	-3	9	12
56	59	-3	9	-2	4	6
57	57	-2	4	-4	16	8
58	60	-1	1	-1	1	1
59	63	0	0	2	4	0
60	64	1	1	3	9	3
61	62	2	4	1	1	2
62	63	3	3	2	4	6
63	63	4	16	2	4	8
531	549	0	~~54~~ 60	0	~~44~~ 52	~~32~~ 46

$\bar{x} = \frac{\Sigma x_i}{n} = \frac{531}{9} = 59$; $\bar{y} = \frac{\Sigma y_i}{n} = 61$

$$r = \cancel{\frac{32}{\sqrt{60 \cdot 44}}} \quad \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

$$= \frac{46}{\sqrt{60 \cdot 52}} = 0.82$$

$\therefore r = 0.82$

Regression line by 'y' on 'x' is

$y = 0.76(x) + 15.76$

'x' on 'y' is

$x = 0.88(y) + 5.03$

$$y - \bar{y} = \frac{\Sigma XY}{\Sigma X^2}(x - \bar{x}) \qquad x - \bar{x} = \frac{\Sigma XY}{\Sigma Y^2}(y - \bar{y})$$

$$y - 61 = \frac{46}{60}(x - 59) \qquad x - 59 = \frac{46}{52}(y - 61)$$

$$y = 0.76(x) + 15.76 \qquad x = 0.88\,y + 5.03$$

~~Karl~~ 4.408

x	y	$X = x - \bar{x}$ $= x - 59.11$	X^2	$Y = y - \bar{y}$ $= y - 120.22$	Y^2	XY
57	113	-2.11	4.45	-7.22		
51	117	-8.11	64.77	-3.22		
62	128	2.89	8.35	7.78		
63	126	3.89	15.13	5.78		
64	130	4.89	23.91	9.78		
65	129	5.89	34.69	8.79		
55	111	-4.11	16.88	-9.22		
58	116	-1.11	1.23	-4.22		
57	112	-2.11	4.45	-8.22		
532	1082	0.01	174.89	0.02	499.55	245.77

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{532}{9} = 59.11$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{1082}{9} = 120.22$$

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 . \Sigma Y^2}} = \frac{245.77}{\sqrt{174.89 . 499.55}}$$

$$r = 0.8314$$

Solve the Karl Pearson's Coefficient of correlation for

x	55	56	57	58	59	60	61	62	63
y	58	59	57	60	63	64	62	63	63

also find Regression line by y on "x" and "x" on "y."

Coefficient of correlation by Regression Line

Equation of regression line 'y' on 'x'.

$$\cancel{y} \bar{y} = \frac{\Sigma XY}{\Sigma X^2} (\cancel{X} - \bar{x})$$

Equation of regression line 'x' on 'y'

$$\cancel{X} - \bar{x} = \frac{\Sigma XY}{\Sigma Y^2} (\cancel{y} - \bar{y})$$

Solve by Karl Pearson's Coefficient of correlation

x	12	14	15	24	20	10	16	24
y	15	25	18	16	20	20	18	25

Also find Regression line By 'Y' on 'X' and 'X' on 'Y'.

31/10/23. Find the coefficent of ~~se~~ corelation and obtain the equation of line of regresion for the given data.

x	6	2	10	4	8	30
y	9	11	5	8	7	40
$X = x - \bar{x}$	0	-4	4	-2	2	0
$Y = y - \bar{y}$	1	3	-3	0	-1	0
XY	0	-12	-12	0	-2	-26
X^2	0	16	16	4	4	40
Y^2	1	9	9	0	1	20

$\bar{x} = \frac{30}{5} = 6$

$\bar{y} = \frac{40}{5} = 8$

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}} = \frac{-26}{\sqrt{40 \cdot 20}}$$

$$r = -0.9192$$

Equation of regression line Y on X:

"y" on "x":

$$y - \bar{y} = \frac{\Sigma XY}{\Sigma X^2}(x - \bar{x})$$

$$y - 8 = \frac{-26}{40}(x - 6)$$

$$\therefore y = -0.65(x) + 11.9$$

$$y = \frac{-26}{40}x + \frac{119}{10}$$

Equation of Regression line

"x" on "y":

$$x - \bar{x} = \frac{\Sigma XY}{\Sigma Y^2}(y - \bar{y})$$

$$x - 6 = \frac{-26}{20}(y - 8)$$

$$\therefore x = -1.3(y) + 16.4$$

$$x = \frac{-26}{20}(y) + \frac{82}{5}$$

H.W

Find the coefficient of ~~corelation~~ correlation and obtain the equation of line of regression for the given data.

(x on y)(y on x)

x	1	2	3	4	5
y	2	5	3	8	7

SPEARMAN'S RANK CORRELATION

A rank correlation coefficient measures the degree of similarity between two rankings and ~~asses~~ access the significance of the relation between them.

For non repeted ranks :

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

where, r = coefficient of rank correlation.
d = difference of ranks
n = number of observations.

For repeated ranks :

$$r = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(l^3 - l) + \frac{1}{12}(m^3 - m) + \frac{1}{12}(n^3 - n) \ldots\right]}{n(n^2-1)}$$

Note: There are three types of problems :

- Ranks are given
- Non repeated ranks
- Repeated ranks.

Two ga in a beauty competition ranked in 12 entries. As follows

x	1	2	3	4	5	6	7	8	9	10	11	12
y	12	7	8	2	3	6	5	4	10	11	9	1

$X = R_1$	$Y = R_2$	$d = R_1 - R_2$	d^2
1	12	-11	121
2	7	-5	25
3	8	-5	25
4	2	2	4
5	3	2	4
6	6	0	0
7	5	2	4
8	4	4	16
9	10	-1	1
10	11	-1	1
11	9	2	4
12	1	11	121
		0	326

$$r = 1 - \frac{6 \sum d^2}{n(n^2-1)} = -0.1398.$$

$\therefore$ Coefficient of rank correlation $r = -0.1398$.

Calculate the coefficient of rank correlation for the following data.

x	2	4	5	6	8	11
y	18	12	10	8	7	5

$n = 6$

x	y	R_1	R_2	$d = R_1 - R_2$	d^2
2	18	6	1	5	25
4	12	5	2	3	9
5	10	4	3	+1	1
6	8	3	4	1	1
8	7	2	5	-3	9
11	5	1	6	-5	25
				0	70

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

$\therefore$ coefficient of rank ~~coefficient~~ correlation $r = -1$.

Find Spearmen's Rank Coefficient for the following data.

x	17	12	22	31	27
y	119	113	117	121	115

$n = 5$

x	y	R_1	R_2	d	d^2
17	119	4	2	2	4
12	113	5	5	0	0
22	117	3	3	0	0
31	121	1	1	0	0
27	115	2	4	-2	4
				0	8

$n = 5$

$\Sigma d^2 = 8$

$$r = 1 - \frac{6 \Sigma d^2}{n(n^2-1)} = 1 - \frac{6(8)}{5(5^2-1)} = 0.6$$

$r = 0.6$

$\therefore$ coefficient of rank correlation

H.W

x	80	92	97	82	64	71	69	58
y	125	135	112	150	109	132	118	120
	4	2	1	3	7	5	6	8
	4	2	7	1	8	3	6	5
	0	0	-6	2	-1	+2	0	3
			36	4	1	4	0	9 /54.

IMP

21/11/2023 Find the Rank correlation for the following data.

x	20	22	28	23	30	30	23	24
y	28	24	24	25	26	27	32	30

$r = 0.247$.

Soln n = no. of observations

$n = 8$

Rank in $x = R_1$

Rank in $y = R_2$

x	y	R_1	R_2	$d = R_1 - R_2$	d^2
20	28	8	3	5	25
22	24	7	7.5	-0.5	+0.25
28	24	3	7.5	-4.5	+20.25
23	25	5.5	6	-0.5	+0.25
30	26	1.5	5	-3.5	+12.25
30	27	1.5	4	-2.5	+6.25.
23	32	5.5	1	4.5	20.25
24	30	4	2	2	4
				-(1)	88.5

$$r = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(l^3 - l) + \frac{1}{2}(m^3 - m) + \dots\right]}{n(n^2 - 1)}$$

In x Series:

23, 30 repeats two times.

$l = 2, m = 2$.

In Y Series:

24 Repeats 2 times.

$n = 2$.

$$r = 1 - \frac{6\left[88.5 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(2^3 - 2)\right]}{8(34 - 1)}$$

$= 1 - 6(0.1785)$

$= 6 - 1.07142$

$= -0.0714$

∴ Coefficient of Rank of rank correlation $r = -0.0714$

From the following table, calculate rank correlation coefficient.

x	45	30	40	9	11	11	60	22	11	52
y	15	15	22	9	15	4	20	9	10	18

n = number of observation = 10.

Rank in $x = R_1$

Rank in $y = R_2$.

$x \rightarrow$ 1 2 3 4 5 6 7 8 9 10

$y \rightarrow$ 1 2 3 4 5 6 7 8 9 10

x	y	R_1	R_2	d	d^2
45	15	3	5	-2	4
30	15	5	5	0	0
40	22	4	1	3	9
9	9	10	8.5	1.5	2.25
11	15	8	5	3	9
11	4	8	10	-2	4
60	20	1	2	-1	1
22	9	6	8.5	-2.5	6.25
11	10	8	7	1	1
52	18	2	3	-1	1
					37.5

Hence, $\Sigma d^2 = 37.5$

In X series: 11 repeats 3 times.

$l = 3$

In Y series: 15, 9 repeats 3, 2 times respectively.

$m = 3 \quad n = 2$

Coefficient of rank correlation:

$$r = 1 - 6\left[\frac{\Sigma d^2 + \frac{1}{12}(l^3 - l) + \frac{1}{12}(m^3 - m) + \frac{1}{12}(n^3 - n)}{n(n^2-1)}\right]$$

$$\frac{37.5 + \frac{1}{12}(3^3-3) + \frac{1}{12}(3^3-3) + \frac{1}{12}(2^3-2)}{10(10^2-1)}$$

$r = 1 - 6\left[0.04\overline{2}\right]$

$= 1 - 0.2545$

$r = 0.7\overline{45}$

$\therefore$ coefficient of rank correlation $r = 0.75$

(H.W) Find the rank correlation for the following data.

i)

x	68	64	75	50	64	80	75	40	55	64
y	62	58	68	45	81	60	68	48	50	70

$d^2 = 72$

$1 - 0.4545$

$r = 0.5454$

ii)

Marks Obtained A	20	30	40	40	50	60	70	80
Marks Obtained B	10	15	20	20	18	25	30	40

$r = 0.9166$

$\Sigma d^2 = 6$

03/11/2023

Random Variables:

Probability:

In a random experiment, let "E" be an event of experiment. Let 'm' be the total number of ~~pes~~ experiment. 'm' is the probable elements in random experiment. 'm' is the probable number of element. Then the probability of E is defined as $p(E)$.

i.e, $p(E) = \frac{m}{n} = \frac{\text{faveriable No. of events}}{\text{Total no. of events}}$

note: • Probability always lies between zero and one.

i.e, $0 \leq p(E) \leq 1$

• Sum of the probabilities is unity of a experiment.

$$\Sigma P_i = 1$$

Random Variable: It is a mapping from sample space (S) to real line ($\mathbb{R}$) i.e $X: S \rightarrow \mathbb{R}$.

Generally, Random Variable are denoted by Capital letters.

Ex: Consider a random experiment consisting of tossing a coin twice. $S = \{HH, HT, TH, TT\}$ is the sample space.

Define a function $X : S \to R$ by selecting number of heads only. i.e, $X(s) = X(HH) = 2$

Similarly, $\Rightarrow X(HT) = 1$

$\Rightarrow X(TH) = 1$

$\Rightarrow X(TT) = 0$.

Types of random Variables: There are two types of random Variables.

- Discrete Random Variable
- Continuous Random Variable.

Discrete Random Variables:

A ~~x~~ random variable which can take only a finite number is called a Discrete Random Variable.

Ex: i) Tossing a coin twice

ii) No. of children in a family.

Continuous Random Variables:

A random variable X is said to be continuous if it can take all possible values in an interval.

Ex: i) In a day through out the temperature.

ii) Hights of the students in a class

Mean (or) Expectation:

Suppose a random Variable (X) assumes the values $x_1, x_2, x_3, \ldots, x_n$ with respect to probability $p_1, p_2, \ldots, p_n$. Then, the expectation (or) mean (or) expected value.

$$\text{Mean } \mu = E(x) = \Sigma p_i x_i$$

Variance: Variance of a random Variable is denoted by σ^2.

$$\sigma^2 = \Sigma p_i x_i^2 - \mu^2$$

$$= E(x^2) - [E(x)]^2$$

Standard Deneation = Sigma.

Probability Destribution Function: $(F(x))$

Let X be a Random Variable. Then the function $F(x)$ is defined by $F(x) = P(X \leq x)$, $-\infty < x < \infty$

It is called the Destribution Function of x.

IMP

A Random Variable X has the following probability function.

x	0	1	2	3	4	5	6	7
$p(x)$	0	k	$2k$	$2k$ ($2k$)	$3k$	k^2	$2k^2$	$7k^2+k$.

Determine i) k

ii) Evaluate $P(X<6)$, $P(X \geq 6)$ and $P(0 < x < 5)$

iii) If $P(X \leq k) > \frac{1}{2}$. find k min.

iv) Determine the destribution function of x.

v) Mean vi) Variance.

We know that Sum of the probabilitis is Unity.

$$\Sigma p_i = 1$$

$$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

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

$k = \frac{1}{10}, -1$ Since, k can't be negative.

$$\therefore k = \frac{1}{10} = 0.1.$$

ii) $P(x < 6) = P(x=1) + P(x=2) + P(x=3) + P(x=4) + P(x=5)$

$$= 0 + k + 2k + 2k + 3k + k^2$$

$$= 8k + k^2.$$

$\because k = \frac{1}{10}$

$$= \frac{8}{10} + \frac{1}{100} = \frac{81}{100}$$

$$= 0.81.$$

$$P(x \geq 6) = P(x=6) + P(x=7)$$

$$= 2k^2 + (7k^2 + k)$$

$\because k = \frac{1}{10}$,

$$= 9k^2 + k$$

$$= 9\left(\frac{1}{100}\right) + \frac{1}{10}$$

$$= \frac{19}{100} = 0.19$$

$$P(0 < x < 5) = P(x=1) + P(x=2) + P(x=3) + P(x=4)$$

$$= k + 2k + 2k + 3k.$$

$\because k = \frac{1}{10}$,

$$= 8k$$

$$= 8\left(\frac{1}{10}\right) = 0.8$$

iii) $P(X \leq k) > 1/2$

If $k = 1 \Rightarrow P(X \leq 1) = P(X = 0) + P(X = 1)$

$= 0 + k$

$= 0.1$

If $k = 2 \Rightarrow P(X \leq 2) = 0 + k + 2k$

$= 3k = 3(1/10)$

$= 0.3$

If $k = 3 \Rightarrow P(X \leq 3) = 0 + k + 2k + 2k$

$= 5k = 5(1/10)$

$= 0.5$

$P(x \leq 3) = 1/2.$

If $k = 4 \Rightarrow P(X \leq 4) = 0 + k + 2k + 2k + 3k$

$= 8(k) = 8(1/10)$

$= 0.8$

$P(X \leq 4) > 1/2 = 0.8.$

$\therefore$ for $P(x \leq k) > 1/2$ 4 is min. value of k.

iv) ~~Mean~~ = Probability distribution function.

x	0	1	2	3	4	5	6	7
$F(x)$	0	k	$3k$	$5k$	$8k$	k^2+8k	$3k^2+8k$	$10k^2+9k$

for $x = 7$

$f(x) = 10k^2 + 9k$

If $k = 0.1$

$= 0.1 + 0.9$

$= 1.$

$\sum p_i x_i = -3(0.05) - 2(0.1) + 0 + 2(0.05) + 2(0.8) + 3(2(0.05))$.
$= 0.8$

$\mu = 0.8$

$\sigma^2 = 9(0.05) + 4(0.1) + 0 + 2(0.05) + 4(0.4) + 9(0.1) - \mu^2$.

$\sigma^2 = 3.5 - (0.8)^2$

$\sigma^2 = 3.5 - 0.64$

$\sigma^2 = 2.86$

$\therefore$ Mean is 0.8 and Variance 2.86 for the given Discrete Destribution.

Find mean and variance of the uniform probability destributi given by $f(x) = \frac{1}{n}$ for $x = 1, 2, 3, \ldots, n$.

x	1	2	3	$\ldots$	n
$p(x)$	$1/n$	$1/n$	$1/n$		$1/n$
$F(x)$					

Mean $= \sum p_i x_i$

$= \frac{1}{n}[1 + 2 + 3 + 4 \ldots n]$

$\frac{1}{\cancel{n}} \frac{\cancel{n}(n+1)}{2}$

$\mu = \frac{n+1}{2}$

Variance $= \sum p_i x_i^2 - \mu^2$

$= \frac{1}{n}[1^2 + 2^2 + 3^2 + \ldots + n^2] - \left(\frac{n+1}{2}\right)^2$

$= \frac{1}{\cancel{n}} \frac{\cancel{n}(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \cancel{\frac{n+1[2n-2]}{6}}$

$n+1\left[\frac{2n+1}{6} - \frac{1}{2}\right]$

$= n+1\left[\frac{2n+1-3}{6}\right]$

$= \frac{(n+1)(2)(n-1)}{6}$

$= \frac{(n+1)(n-1)}{3}$

$$= \frac{n+1}{2}\left[\frac{2n+1}{3} - \frac{n+1}{2}\right]$$

$$= \frac{n+1}{2}\left[\frac{4n+2-3n-3}{6}\right]$$

$$= \frac{(n+1)(n-1)}{12}$$

$$\sigma^2 = \frac{n^2-1}{12}.$$

$\therefore$ Mean $= \frac{n+1}{2}$, Variance $= \frac{n^2-1}{12}$

IMP

Let X denote the <u>maximum</u> of two numbers. Two dice are thrown. Let X assign to each point (A, B) in the maximum of its numbers.

i.e, $X(a,b) = \max(a,b)$.

Find the probability destribution. X is a random variable with $X(S) = \{1,2,3,4,5,6\}$. Also find the mean and variance of the destribution.

04/11/23

The total number of cases are $6 \times 6 = 36$. The maximum number would be 1, 2, 3, 4, 5, 6.

i.e, $X(S) = X(a,b) = \max(a,b)$

The number on will appear only in one case

$(1,1) \Rightarrow P(1) = P(X=1) = \frac{1}{36}$

$(1,2)(2,1)(2,2) \Rightarrow P(2) = P(X=2) = \frac{3}{36}.$

$$S = \left\{ \begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix} \right\}$$

for max 3: $P(3) = P(X=3) = \frac{5}{36}$ (3,1)(3,2)(3,3)(2,3)(1,3)

for max 4: $P(4) = P(X=4) = \frac{7}{36}$ (4,1)(4,2)(4,3)(4,4)(1,4) (2,4)(3,4)

for max 5: $P(5) = P(X=5) = \frac{9}{36}$ (5,1)(5,2)(5,3)(5,4)(5,5) (1,5)(2,5)(3,5)(4,5)

for max 6: $P(6) = P(X=6) = \frac{11}{36}$. (6,1)(6,2)(6,3)(6,4)(6,5) (6,6)(1,6)(2,6)(3,6)(4,6) (5,6).

Required Probability Distribution :

x	1	2	3	4	5	6
$P(x)$	$1/36$	$3/36$	$5/36$	$7/36$	$9/36$	$11/36$

$$\text{Mean} = \mu = \sum p_i x_i$$

$$= \frac{1}{36} + \frac{6}{36} + \frac{15}{36} + \frac{28}{36} + \frac{45}{36} + \frac{66}{36}$$

$$\mu = 4.47$$

$$\sigma^2 = \sum p_i x_i^2 - \mu^2$$

$$= \frac{1}{36} + \frac{12}{36} + \frac{45}{36} + \frac{16\times7}{36} + \frac{81}{36} + \frac{36\times11}{36} - 4.47^2.$$

$$= 21.97 - 4.47^2 = 1.989$$

(HW) Let X denote the minimum of two numbers that appear when a pair of fair dies is thrown once. Determine

i) Discrete Probability Distribution.

ii) Expectation (mean)

iii) Variance

A sample of 4 items are selected at random from a box containing 12 items of which 5 are defective. Find the expected number E of defective items.

34 495

Let X denote the number of defective items! 175

among 4 items drown from 12 items 210

70

X can take the values 0,1,2,3,4. 5

The number of defective items = 5

The no. of ~~doo~~ non-defective items = 7.

$P(X=0) = P(\text{no. of defective} = 0) = \frac{{}^7C_4 \, {}^5C_0}{{}^{12}C_4} = \frac{35}{495} \quad \frac{7}{99}$

$$\boxed{{}^nC_r = \frac{n!}{r!\,(n-r)!}}$$

${}^nC_1 = n \quad {}^nC_n = 1$

${}^nC_0 = 1$

$P(X=1) = P(\text{one defect}) = \frac{{}^7C_3 \, {}^5C_1}{{}^{12}C_4} = \frac{175}{495} \quad \frac{35}{99}$

$P(X=2) = P(\text{two defect}) = \frac{{}^7C_2 \, {}^5C_2}{{}^{12}C_4} = \frac{210}{495} = \frac{14}{33}$

$P(X=3) = P(\text{three defect}) = \frac{{}^7C_1 \, {}^5C_3}{{}^{12}C_4} = \frac{70}{495} \quad \frac{14}{99}$

$P(X=4) = P(\text{all defects}) = \dfrac{{}^{7}C_0 \, {}^{5}C_4}{{}^{12}C_4} = \dfrac{5}{495} = \dfrac{1}{99}$

$1.\overline{666}$

$3.\overline{48}$

$0.7\overline{070}$

Required Probability Destribution.

X	0	1	2	3	4
P(x)	$\frac{7}{99}$	$\frac{35}{99}$	$\frac{42}{99}$	$\frac{14}{99}$	$\frac{1}{99}$.

Mean (Expectation) $= \sum_{i=0}^{4} p_i x_i$

$$\mu = 0 + \frac{35}{99} + \frac{84}{99} + \frac{42}{99} + \frac{4}{99}$$

$$\mu = 1.\overline{6}$$

Variance $= \sum_{i=0}^{4} p_i x_i^2 - \mu^2$

$$= 0 + \frac{35}{99} + \frac{4 \times 42}{99} + \frac{9 \times 14}{99} + \frac{16}{99} - (1.\overline{6})^2$$

$$= 3.\overline{48} - 1.\overline{6}^2 = 0.\overline{70}$$

From a lot of 10 items containing 3 defective, a sample of 4 items is drawn at random. Let the random Variable X be no. of defective items. Find the probabality destribution of X. when the sample is drown without replacement.

A ~~Pair~~ player tosses 3 fair coines. He wins ₹500 if 3 heads appeare. He wins ₹300 if 2 heads appeare. On the other hand, He lose ₹1500 If 3 Tails accure. Find the expected gain of the player.
+₹100 if one head appears.

Let X denote the gain. Then the range $x = \{-1500, 100, 300, 500\}$

The sample space $S = \{HHH, HTH, HTT, HHT, THH, THT, TTH, TTT\}$

The probability of getting all three heads: (Getting ₹500)

$P(X=3) = \frac{1}{8}$

The probability of getting 2 heads: $P(X=2) = \frac{3}{8}$ (Getting ₹300)

The probability of getting 1 head $P(X=1) = \frac{3}{8}$ (Getting ₹100)

The probability of getting no head (lossing ₹1500)

$P(X=0) = \frac{1}{8}$.

Required Descrete Probability Destribution

x	-1500	100	300	500
$P(x)$	$1/8$	$3/8$	$3/8$	$1/8$

$E(x) = \Sigma P_i \cdot x_i$

$= +1500\left(\frac{-1}{8}\right) + 100(3/8) + 300\left(\frac{3}{8}\right) + 500\left(\frac{1}{8}\right)$

$= 25$

A player wins and gets 5 on a single throw of a dice.
He losses if he gets 2 or 4. If he wins He gets ₹ 50.
If he lose He gets ₹ 10. other wise he must pay ₹ 15.

Let X denote the gain.

The range $X = \{-15, 10, 50\}$

The sample space $S = \{1, 2, 3, 4, 5, 6\}$

The probability of getting ₹ 50 $= P(5) = \frac{1}{6}$.

The probability of getting ₹ 10 $= P(2) + P(4) = \frac{1}{6} + \frac{1}{6}$

$= \frac{2}{6}$.

The probability of lossing ₹ 15 $= P(1) + P(3) + P(6)$

$= \frac{3}{6}$.

Required Discrete Probability Distribution -

x	-15	10	50
$P(x)$	$3/6$	$2/6$	$1/6$

$E(x) = \Sigma p_i x_i$

$= -15\left(\frac{3}{6}\right) + 10\left(\frac{2}{6}\right) + 50\left(\frac{1}{6}\right)$

$\mu = 4.1\bar{6}$

Game is favouriable to player $\because E(x) > 0$.

Probability Density function or (Probability Mass)

Continuous Probability Distribution :

Properties of Probability Destribution :

The density function $f(x)$ satisfies

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

ii) $f(x) \geq 0 \quad \forall\, x \in R$

Mean:

Mean of a destribution is given by

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

Variance:

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

Median:

Median is the point which devides the entire destributi into two equal parts. Thus, If x is defined from a to b. and "m" is median i.e,

$$\text{Median} = \int_a^m f(x)\cdot dx = \int_m^b f(x)\,dx = \frac{1}{2}.$$

Mode: Mode is the value of x for which $f(x)$ is Max. Mode is given by $f'(x) = 0$, $f''(x) < 0$ for $a < x < b$.

If a random variable has the probability density

$$f(x) = \begin{cases} 2e^{-2x}, & \text{for } x > 0 \\ 0, & \text{for } x \leq 0 \end{cases}$$ find the probabilities

what it'll take on a value (i) between $(1,3)$, $P(1 \leq x \leq 3)$

(ii) > 0.5, $P(X > 0.5)$

(i) $P(1 \leq x \leq 3)$

$$= \int_1^3 f(x)\,dx$$

$$= \int_1^3 2e^{-2x}\,dx \qquad \left[\because e^{-ax} = \frac{e^{-ax}}{-a}\right]$$

$$= 2\int_1^3 e^{-2x}\,dx$$

$$= 2\left[\frac{e^{-2x}}{-2}\right]_1^3$$

$$= -\left[e^{-2(3)} - e^{-2(1)}\right]$$

$$= e^{-2} - e^{-6}.$$

ii) $P(X \geq 0.5)$

$$= \int_{0.5}^{\infty} f(x)\,dx = \int_{0.5}^{\infty} 2e^{-2x}\,dx$$

$$= 2\int_{0.5}^{\infty} e^{-2x}\,dx$$

$$= \frac{2}{-2}\left[e^{-2x}\right]_{0.5}^{\infty} = -\left[e^{-2(\infty)} - e^{-2(0.5)}\right]$$

$$= e^{-1} = 1/e.$$

9/11/23

The probability function $f(x)$ continuous random variable is given by $f(x) = ce^{-|x|}$. Show that $c = \frac{1}{2}$. Find the mean and ~~variantion~~ variance of the distribution. Also find Probability that the variat lies between 0 and 4.

Given: $f(x) : ce^{-|x|}$, $-\infty < x < \infty$

we know that sum of probability is unity.

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

$[\because c$ is constant $]$

$\Rightarrow \int_{-\infty}^{\infty} ce^{-|x|}\,dx = 1$ $\quad \left[\because \int_{-\infty}^{\infty} f(x)dx = 2\int_{0}^{\infty} f(x)dx\right]$

$\Rightarrow 2c\int_{0}^{\infty} e^{-|x|}\,dx = 1$ $\quad [|x| = x]$

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

$\Rightarrow 2c\left[\frac{e^{-x}}{-1}\right]_0^{\infty} = 1 \Rightarrow -2c\left[e^{-x}\right]_0^{\infty} = 1$

$\Rightarrow -2c\left[e^{-\infty} - e^{-0}\right] = 1$

$\Rightarrow -2c\,[0-1]$

$\Rightarrow 2c = 1$

$\Rightarrow \boxed{c = \frac{1}{2}}$

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

$\mu = \int_{-\infty}^{\infty} x\left[ce^{-|x|}\right]dx$

$$\mu = c\int_{-\infty}^{\infty} x e^{-|x|} dx \qquad \because c = 1/2$$

Let x is odd and e^{x} be the even function.

$$\Rightarrow \mu = \frac{1}{2}\int_{-\infty}^{\infty} x e^{-|x|} = 0$$

$\mu = 0.$

Variance: $\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x) - \mu^2$

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 \left[c\ e^{-|x|}\right] dx - \mu^2$$

$$\left[c = \frac{1}{2}, \mu = 0\right]$$

$$\sigma^2 = \frac{1}{2}\int_{-\infty}^{\infty} x^2 e^{-x} dx$$

$$= \frac{1}{2}\, 2\int_{0}^{\infty} x^2 e^{-x} dx$$

$$= \int_{0}^{\infty} x^2 e^{-x} dx$$

$$= \left[x^2 \frac{d}{dx}(e^{-x}) - 2x \int e^{-x} dx + 2\frac{d}{dx}(e^{-x})\right]_0^{\infty}$$

$$= \left[x^2\left(\frac{e^{-x}}{-1}\right) - 2x\left(\frac{e^{-x}}{-1}\right) + 2\left(\frac{e^{-x}}{-1}\right)\right]_0^{\infty}$$

$$= \cancel{x^2 - 2x + 2} \quad 0 - \left[0 - 0 + 2\left(\frac{e^{0}}{-1}\right)\right]$$

$$\sigma^2 = -[-2]$$

$$\sigma^2 = 2$$

Consider $P(0 \le x \le 4)$

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^4 c\, e^{-|x|}\,dx$$

$$= \frac{1}{2}\int_0^4 e^{-x}\,dx$$

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

$$= -\frac{1}{2}\left[e^{-4} - e^{0}\right]$$

$$= \frac{1-e^{-4}}{2}$$

$$= 0.49$$

A continuous random Variable x has he probability density function $f(x) = \begin{cases} kx\, e^{-\lambda x}, & \text{for } x \geq 0,\ \lambda > 0. \\ 0, & \text{other wise} \end{cases}$

(i) Find k. (ii) Mean (iii) Variance.

Sum of the probability is unity.

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

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

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

$$k\int_0^{\infty} x\,e^{-\lambda x}\,dx = 1$$

$$k\left[x\,\frac{d}{dx}(e^{-\lambda dx}) - (1)\int e^{-\lambda dx}\,dx\right]_0^{\infty} = 1$$

$$k\left[0 - \left[-\frac{e^{0}}{\lambda^2}\right]\right] = 1 \Rightarrow \boxed{k = \lambda^2}$$

Mean:

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

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

$$= 0 + \int_0^{\infty} x\left[kx\,e^{-\lambda x}\right]dx$$

$$= \lambda^2\int_0^{\infty} x^2\,e^{-\lambda x}\,dx$$

$$= \lambda^2\left[x^2\left(\frac{e^{-\lambda x}}{-\lambda}\right) - 2x\left(\frac{e^{-\lambda x}}{+\lambda^2}\right) + 2\left(\frac{e^{-\lambda x}}{-\lambda^3}\right)\right]_0^{\infty}$$

$$= \lambda^2\left[0 - \left[0 - 0 + 2\left(\frac{1}{-\lambda^3}\right)\right]\right]$$

$$= \lambda^2\left(2/\lambda^3\right)$$

$\mu = \frac{2}{\lambda}$.

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

$$\sigma^2 = 0 + \int_0^{\infty} x^2 \left[kx\, e^{-\lambda x} \right] dx - \left(\frac{2}{\lambda}\right)^2$$

$$= \lambda^2 \int_0^{\infty} x^3 e^{-\lambda x}\, dx \qquad \text{(Bernolis)}$$

$$= \lambda^2 \left[+x^3 \left(\frac{e^{-\lambda x}}{-\lambda}\right) - 3x^2 \left(\frac{e^{-\lambda x}}{+\lambda^2}\right) + 6x \left(\frac{e^{-\lambda x}}{-\lambda^3}\right) - 6 \left(\frac{e^{-\lambda x}}{+\lambda^4}\right)\right]_0^{\infty} - \frac{4}{\lambda^2}$$

$$= \lambda^2 \left[0 - \left(0 - 0 + 0 - 6 \left(\frac{e^0}{+\lambda^4}\right)\right)\right] - \frac{4}{\lambda^2}$$

$$\lambda^2 \left[+ \frac{6}{\lambda^4} \right] - \frac{4}{\lambda^2}$$

$$\sigma^2 = + \frac{6}{\lambda^2} - \frac{4}{\lambda^2}$$

$$= \frac{2}{\lambda^2}$$

H.W

The continues for cont. Probability function

$f(x) = kx^2 e^{-x}$ $(x \geq 0)$.

find k, mean, variance.

The probability Density function of a random variable x is $f(x) = \begin{cases} \frac{1}{2} \sin x, & 0 \leq x \leq \pi \\ 0, & \text{else where.} \end{cases}$

Find Mean, Mode, Variance and also find probability between 0 and $\pi/2$.

10/11/23

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

$$= \int_{-\infty}^{0} x f(x) dx + \int_{0}^{\pi} x f(x) dx + \int_{\pi}^{\infty} x f(x) dx$$

$$\because f(x) = \begin{cases} \frac{1}{2} \sin x, & 0 \le x \le \pi \\ 0, & \text{else} \end{cases}$$

$$\Rightarrow \quad \mu = 0 + \int_{0}^{\pi} x \frac{1}{2} \sin x \, dx + 0$$

$$= \frac{1}{2} \int_{0}^{\pi} x \sin x \, dx$$

$$= \frac{1}{2} \left[x(-\cos x) - 1(-\sin x) \right]_0^{\pi}$$

$$= \frac{1}{2} \left[-x\cos x + \sin x \right]_0^{\pi}$$

$$= \frac{1}{2} (-\pi(-1) + 0) - (0)$$

$$\boxed{\mu = \frac{\pi}{2}}$$

Mode:

Given: $f(x) = \frac{1}{2} \sin x$.

$f(x)$ to be maximum, $f'(x) = 0$.

$$f'(x) = \frac{1}{2} \cos x = 0$$

$$\Rightarrow x = \frac{\pi}{2}.$$

$f''(x) = \frac{-1}{2} \sin x$ at $x = \frac{\pi}{2}$.

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

$f''(x) = \frac{-1}{2} < 0$

$f(x) = \frac{1}{2} \sin x$

$f'(x) = \frac{1}{2} \cos x$

at $x = \pi/2$ it is Max.

$f''(x) = -\frac{1}{2}$ at $x = \frac{\pi}{2}$

Hence, $\boxed{\text{Mode} = \pi/2}$.

Median : It is the value of x which divides the entire destribution into two equal parts.

Hence, $$\int_a^m f(x)\,dx = \int_m^b f(x)\,dx = 1/2$$

$a = 0, \; b = \pi$

$$\int_0^m \frac{1}{2} \sin x \, dx = \frac{1}{2}$$

$$\int_0^m \sin x \, dx = 1$$

$$\left[-\cos x\right]_0^m = 1$$

~~$-\cos m$~~

$-\cos m + \cos 0 = 1$

$-\cos m + 1 = 1$

$\cos m = 0$

$\Rightarrow m = \cos^{-1}(0)$

$m = \frac{\pi}{2}$

$$\boxed{\therefore \text{Median} = \frac{\pi}{2}}$$

(24)

If X is a continuous random variable and $Y = aX + b$. Prove that $E(Y) = a\,E(X) + b$ and $V(Y) = a^2 V(X)$ where 'V' stands for Variance. and 'a', 'b' are constants.

By the definition of mean, $E(x) = \int_{-\infty}^{\infty} x\, f(x)\, dx$.

Since, $Y = aX + b$ — ①

$\Rightarrow$ ~~E(Y)~~ $E(Y) = E(aX + b)$

$$= \int_{-\infty}^{\infty} (aX + b)\, f(x)\, dx$$

$$= a\int_{-\infty}^{\infty} x\, f(x)\, dx + b\int_{-\infty}^{\infty} f(x)\, dx$$

$$E(Y) = a\, E(x) + b\,(1)$$

[$\because$ Sum of all probability is unity, $\int_{-\infty}^{\infty} f(x)\, dx = 1$]

$$E(Y) = a\, E(x) + b \quad — ②$$

① - ② $\Rightarrow$ $[Y - E(Y)] = ax + b - [a\,E(X) + b]$

$$[Y - E(Y)] = a\,[X - E(X)]$$

Squaring on both sides

$$[Y - E(Y)]^2 = a^2\,[X - E(X)]^2$$

$$E\,[Y - E(Y)]^2 = a^2\, E\,[X - E(X)]^2$$

$$V(Y) = a^2\, V(X).$$

A continuous Random variable is defined by

(H.W)

$$f(x) = \begin{cases} \frac{1}{16}(3+x)^2, & \text{if } -3 \le x \le -1 \\ \frac{1}{16}(6-2x^2), & -1 \le x \le +1 \\ \frac{1}{16}(3-x)^2, & +1 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$$

Verify that $f(x)$ is a density function and also find the mean of x.

$x = \frac{1}{10}$ $\mu = 0$.

(H.W)

If x is the continuous random variable whose density function is

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

Find => $E[25x^2 + 30x - 5]$

$$25\,E(x^2) + 30\,E(x) - 5$$

$$\int x^2 f(x)\,dx \qquad \int x f(x)\,dx$$

BINOMIAL DISTRIBUTION:

A random variable x has a binomial distribution if it assumes only non negative values and its probability distribution is given by

$$P(X=r) = p(r) = \begin{cases} {}^nC_r \, p^r q^{n-r}, & r = 0,1,2,\dots n \\ & q = 1-p \\ 0, & \text{other wise} \end{cases}$$

Mean of Binomial Distribution:

The Binomial Distribution is given by

$$p(r) = {}^nC_r \, p^r q^{n-r}, \quad r = 0,1,2,\dots,n, \quad q = 1-p$$

Mean of X: $\mu = E(x) = \sum_{r=0}^{n} r \, p(r)$

$$\mu = \sum_{r=0}^{n} r \left({}^nC_r \, p^r q^{n-r}\right)$$

$$\mu = {}^nC_1 p^1 q^{n-1} + 2\,{}^nC_2 p^2 q^{n-2} + \dots + (n)\,{}^nC_n \, p^n q^{n-n}$$

$$= np^1 q^{n-1} + 2 \frac{n(n-1)}{2!} p^2 q^{n-2} + \dots + n(1) p^n q^0$$

$$= npq^{n-1} + n(n-1) p^2 q^{n-2} + \frac{n(n-1)(n-2)}{2!} p^3 q^{n-3} + np^n$$

$$= np\left[q^{n-1} + (n-1) p^1 q^{n-2} + \frac{(n-1)(n-2)}{2!} p^2 q^{n-3} + p^{n-1}\right]$$

$$= np\left[q+p\right]^{n-1}$$

(according to Binomial Destribution)

$$\left[\because q = 1-p\right]$$

$$= np\,[1]^{n-1}$$

$\mu = np$ $\therefore$ Mean of Binomial Destribution is np.

Variance of the Binomial Destribution:

Since, $p(r) = {}^nC_r\, p^r q^{n-r}$, $r = 0, 1, \ldots, n$, $q = 1 - p$.

$$\sigma^2 = \sum_{r=0}^{n} r^2 p(r) - \mu^2$$

$$\sigma^2 = \sum_{r=0}^{n} \cancel{r^2 + (r - r)}(r^2 + r - r)\, p(r) - \mu^2$$

$$= \sum_{r=0}^{n} [r(r-1) + r]\, p(r) - (np)^2$$

$$= \sum_{r=0}^{n} r(r-1)\, p(r) + \sum_{r=0}^{n} r\,p(r) - (np)^2$$

$$= \left[\sum_{r=0}^{n} r(r-1)\, {}^nC_r\, p^r q^{n-r}\right] + np - n^2p^2$$

$$= \Big[0 + 1(0) + 2(2-1)\, {}^nC_2\, p^2 q^{n-2} + 3(3-1)\, {}^nC_3\, p^3 q^{n-3}$$

$$+ n(n-1)\, {}^nC_n\, p^n q^{n-n}\Big] + np - n^2p^2.$$

$$= \Big[2\,\frac{n(n-1)}{2!}\, p^2 q^{n-2} + 6\,\frac{n(n-1)(n-2)}{3!} + \ldots + n(n-1)\, p^n\Big] + np - n^2p^2$$

Recurrence Relation for Binomial Distribution:

we know that $p(r) = {}^nC_r\, p^r q^{n-r}$ — ①

Then, $p(r+1) = {}^nC_{r+1}\, p^{r+1} q^{n-r-1}$ — ②

$$\text{②}/\text{①} \Rightarrow \frac{p(r+1)}{p(r+1)} = \frac{{}^nC_r\, p^r\, q^{n-r}}{{}^nC_{r+1}\, p^{r+1}\, q^{n-r-1}}$$

$$= \frac{\dfrac{\cancel{n!}}{(r+1)!(n-r-1)!}}{\dfrac{\cancel{n!}}{r!\,(n-r)!}}\; p^{r+1-r}\; q^{n-r-n+r+1}$$

$$= \frac{\cancel{r!}\,(n-r)\,\cancel{(n-r-1)!}}{(r+1)\,\cancel{r!}\,\cancel{(n-r-1)!}}\; p^{+1} q^{-1}$$

$$\frac{p(r+1)}{p(r)} = \frac{n-r}{r+1} \cdot \frac{p}{q}$$

$$\boxed{p(r+1) = \frac{(n-r)\,p}{(r+1)\,q} \cdot p(r).}$$

11/17/23

A fair coin is tossed 6 times Find the probability of getting 4 Heads.

$n = 6$

$p = \frac{1}{2}$, $q = \frac{1}{2}$.

$p(r) = {}^nC_r\, p^r q^{n-r}$

Probability of getting head $p = \frac{1}{2}$.

$$p(4) = {}^6C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^{6-4}$$

$$= \frac{6 \times 5 \times 4!}{4! \times 2!} \quad \frac{1}{2^6}$$

$$P(X=4) = \frac{15}{64} = 0.234375$$

∴ Probability of getting 4 Heads is $\frac{15}{64}$.

Ten coins are thrown simultaneously. Find the probability of getting atleast i) 7 Heads $X \geq 7$

ii) 6 Heads $X \geq 6$

iii) 1 Head. $X \geq 1$

for $P(r) = {}^nC_r\, p^r q^{n-r}$

$n = 10,\ p = \frac{1}{2},\ q = \frac{1}{2}$.

i) Probability of getting atleast 7 heads:

$$P(X \geq 7) = P(X=7) + P(X=8) + P(X=9) + P(X=10)$$

$$= {}^{10}C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{10-7} + {}^{10}C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{10-8}$$

$$+ {}^{10}C_9 \left(\frac{1}{2}\right)^9 \left(\frac{1}{2}\right)^{10-9} + {}^{10}C_{10} \left(\frac{1}{2}\right)^{10} \left(\frac{1}{2}\right)^{10-10}.$$

$$= \frac{15}{128} + \frac{45}{1024} + \frac{5}{512} + \frac{1}{1024}$$

$$= \frac{120+45+10+1}{1024} = \frac{176}{1024} = 0.1718$$

ii) $P(X \geq 6) = P(X=6) + P(X=7) + P(X=8) + P(X=9) + P(X=10)$ 27

$= P(X=6) + P(X \geq 7)$

$= {}^{10}C_6 \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^{10-6} + \frac{176}{1024}$

$\frac{105}{512} + \frac{176}{1024}$

$= \frac{210 + 176}{1024} = \frac{386}{1024} = 0.37695$

iii) Probability of getting atleast one head:

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

$= 1 - {}^{10}C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{10-0}$

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

$= \frac{1023}{1024} = 0.999.$

H.W

Determine the probability of getting the sum 6. exactly 3 times in 7 throws with pair of fair dice.

The binomial destribution for which the mean is 4 and Variance is 3.

Given: $\mu = np = 4.$ — ①

$\sigma^2 = npq = 3$ — ②

$②/① = \frac{\cancel{np}q}{\cancel{np}} = \frac{3}{4} \Rightarrow \boxed{q = 3/4}$

Since, $p = 1 - q$

$\Rightarrow p = 1 - 3/4$

$\boxed{p = 1/4}$

(1) $\Rightarrow$ $npq = 3$

$n\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = 3$

$\boxed{n = 16}$, $p = \frac{1}{4}$, $q = \frac{3}{4}$.

$\because P(r) = {}^nC_r \, p^r q^{n-r}$.

Hence Binomial Destribution

$$P(r) = \begin{cases} {}^{16}C_r \left(\frac{1}{4}\right)^r \left(\frac{3}{4}\right)^{16-r} & , r = 0, 1, 2, \dots 16 \\ 0 & , \text{otherwise} \end{cases}$$

The mean and variation of a Binomial Destribution is 2, 8/5.

Find n.

Given: $np = 2$

$npq = \frac{8}{5}$

(2) / (1) $= \frac{\cancel{np}q}{\cancel{np}} = \frac{8/5}{2}$

$\Rightarrow \boxed{q = \frac{4}{5}}$ $\Rightarrow p = 1 - q$

$= 1 - \frac{4}{5}$

$\boxed{p = \frac{1}{5}}$

(2) $\Rightarrow$ $n\left(\frac{1}{5}\right)\left(\frac{4}{\cancel{5}}\right) = \frac{8}{\cancel{5}}$

$n = \frac{8}{4} \times 5 = 10.$ $\boxed{\therefore n = 10}$

(H.W) ... probability of getting the sum 6 exactly 3 times in 7 throws with pair of fair dice.

In a single throw of a pair of fair dice, Sum of 6 can occur in 5 ways i.e, $\{(1,5), (5,1), (4,2), (2,4), (3,3)\}$ out of $6^2 = 36$.

P = probability of occurence of sum 6 in one throw $= \frac{5}{36}$.

$\Rightarrow q = 1 - p = \frac{31}{36}$.

Given : $n = 7$ (7 throws).

$\therefore$ Probability of getting 6 exactly 3 times in 7 throws $= P(X = 3)$

$\because P(X = r) = {}^nC_r \, p^r q^{n-r}$

$P(X = 3) = {}^7C_3 \, p^3 q^{7-3}$

$= {}^7C_3 \left(\frac{5}{36}\right)^3 \left(\frac{31}{36}\right)^{7-3}$

$= 0.051559$

(H.W)

A dice is thrown 6 times. If getting an even no. is a success. Find Probability of atleast on success.
≤ 3 succes, 4 Succes.

The incidence of an occupational disease in and industry is such that the workers have a 20%. chance of suffering from it. what is the probability that out of 6 workers choisen at random, 4 or more will suffer from desise (Virus)

The probability of worker suffering from disease $p = 20\% = \frac{1}{5} = 0.2$

$${}^6C_4 \left(\frac{1}{5}\right)^4 \left(\frac{4}{5}\right)^{6-4} = \frac{48}{3125}$$

$\because q = 1 - p = \frac{4}{5} = 0.8$

$n = 6$

Probability of 4 or more suffring $= P(X \geq 4)$

$\Rightarrow P(X = r) = {}^nC_r\, p^r q^{n-r}$

$$P(X = 4) = {}^6C_4 (0.2)^4 (0.8)^{6-4}$$

$$= \frac{48}{3125} = 0.01536$$

$$P(X = 5) = {}^6C_5 (0.2)^5 (0.8)^{6-5}$$

$$= \frac{24}{15625} = 0.001536$$

$$P(X = 6) = {}^6C_6 (0.2)^5 (0.8)^{6-6}$$

$$= \frac{1}{15625} = 0.000064$$

$$P(X \geq 4) = P(X = 4) + P(X = 5) + P(X = 6)$$

$$= \frac{48}{3125} + \frac{24}{15625} + \frac{1}{15625}$$

$$= \frac{240 + 24 + 1}{1562[illegible]} = \frac{265}{[illegible]} = 0.0169$$

If 3/20 are defective and 4 are chosen randomly for inspection. what is the Probability that only one of the defective type will be included.

Let P : Probability of defective type :

$p = 3/20$ (Given)

$n = 4$ (Given)

$q = \frac{17}{20}$ $(\because q = 1-p)$

To find : $P(X=1)$ = exactly one will be defective.

$P(X=r) = {}^nC_r \, p^r \, q^{n-r}$

$P(X=1) = {}^4C_1 \left(\frac{3}{20}\right)^1 \left(\frac{17}{20}\right)^{4-1}$

$= 0.368475$ 19.023756

IMP) Out of 800 families with 5 children each, how many would you expect to have

a) 3 boys
b) 5 girls
c) 2 or 3 boys
d) atleast one boy

Assume equal probabilities for boys and girls.

let probability of no. of boys in each family be $p = \frac{1}{2}$, of girl $q = \frac{1}{2}$.

no. of boys in family = X.

no. of children => $n = 5$.

i) Probability of having 3 boys :

$P(X = 3)$.

$$P(X=3) = {}^{5}C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{5-3}$$

$$= \frac{5}{16}$$

For 800 families, the ~~pro~~ probability of no. of families having 3 boys:

$$\Rightarrow \frac{5}{16} \times 800 = 250 \text{ families.}$$

ii) 5 girls

$$P(X=0) = {}^{5}C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{5-0}$$

$$= \frac{1}{32}.$$

For 800 famils, the no. of families having 5 girls

i.e, $\frac{1}{32} \times 800 = 25$

iii) 2 or 3 boys

$$= P(X=2) + P(X=3)$$

$$= {}^{5}C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{5-2} + {}^{5}C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{5-3}$$

$$= \frac{5}{16} + \frac{5}{16}$$

$$= \frac{10}{16} = \frac{5}{8}$$

For 800 families, the probability of no. of families having 2 or 3 boys is $\frac{10}{16} \times 800$

= 500 families

iv) Atleast one boy

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

$$= 1 - \frac{1}{32}$$

$$= \frac{31}{32}.$$

For 800 families the probability of no. of families having atleast one boy is

$$\frac{31}{32} \times 800 = 775.$$

16/11/23 In 8 throws of a dice, 5 or 6 is considered as success. Find the mean number of success, Standard Deviation. (σ)

p = Probability of Success. $= 2/6 = 1/3$

$\therefore q = 1 - p \Rightarrow q = 1 - \frac{1}{3}$

$$q = \frac{2}{3}$$

Given: $n = 8$.

$\therefore$ Mean $= np$

$$\mu = 8 \times \frac{1}{3} = \frac{8}{3} = 2.66..$$

Variance $= \sigma^2 = npq = 8 \times \frac{1}{3} \times \frac{2}{3} = \frac{16}{9}$

Standard deviation $= \sigma = \sqrt{npq} = \frac{4}{3} = 1.33.$

$\therefore$ Mean is 2.66 and Standard Deviation is 1.33

If the probability of defictive bolt is 0.2. Find mean, standrad deveation for the destribution of bolt's in a total of 400.

Given:- $p = 0.2$

$n = 400$

$\because q = 1 - p$

$q = 1 - 0.2$

$q = 0.8$

$\because$ Mean $= np = 400(0.2)$

$\mu = 80$

$\because$ Variance $= npq = 400(0.2)(0.8)$

$\sigma^2 = 16$

$\because$ Standard Deveatio $= \sqrt{npq}$

$\sigma = \cancel{\sqrt{64}}\ \sqrt{64}$

$\sigma = \cancel{8}\ 8$

$\therefore$ Mean is 80 and Standard deveation is

Fitting a Binomial Destribution :

Fit a binomial destribution to the following data

	x	0	1	2	3	4	5
Observed Frequency	$f(x)$	2	14	20	34	22	8

Also find the expected frequency.

Given : $n = 6$

Total frequency $\Sigma f_i = 2 + 14 + 20 + 34 + 22 + 8$

$N = \Sigma f_i = 100$

Mean $= \Sigma f_i x_i / \Sigma f_i$

$= \frac{0(2) + 1(14) + 2(20) + 3(34) + 4(22) + 5(8)}{100}$

$= \frac{0 + 14 + 40 + 102 + 88 + 40}{100}$

$= \frac{284}{100}$

$\mu = 2.84$

For fitting a binomial distribution, take mean of binomial distribution = mean of given data.

i.e, $np = \mu$

$5p = 2.84$

$\Rightarrow p = 0.57$

$\Rightarrow q = 0.43.$

$6p = 2.84$

$p = 0.47\bar{3}$ (7/15)

$q = 0.52\bar{6}$ (79/150)

Binomial Distribution of random variable X is

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

$= {}^5C_x (0.57)^x (0.43)^{5-x}$ — (1)

$X = x$	$P(X=x) = {}^5C_x (0.57)^x (0.43)^{5-x}$		~~${}^6C_x (0.473)^x (0.526)^{6-x}$~~	Expected freq. $N P(x)$.
$X = 0$	0.0147	(or)	~~0.0213~~	1.47
$X = 1$	0.0974		~~0.1150~~	9.74
$X = 2$	0.2583		~~0.2585~~	25.83
$X = 3$	0.3424		~~0.3098~~	34.24
$X = 4$	0.2269		~~0.2088~~	22.69
$X = 5$	[illegible]		~~0.0750~~	6.02
				99.99 [illegible]

- Expected frequence are close to absolute frequencies.
- Expected frequencies can be rounded off to the nearest integer to get expected frequencies as while numbers.

Expected frequency

1, 10, 26, 34, 23, 6

x	0	1	2	3	4	5
$f(x)$	2	14	20	34	22	8
Exp. $f(x)$	1	10	26	34	23	6

Test Next week
4 Q
write 2 Q

All are IMP

C.W

Fit a binomial destribution to the folloing frequency destribution.

x	0	1	2	3	4	5	6
$f(x)$	13	25	52	58	32	16	4

$N = 200$

$u = \frac{535}{200} = 2.675$

$p = 0.4458 =$

$q = \frac{133}{240}$ 0.

$x=0$	0.028962	5.7944
$x=1$	0.139806	27.9612
$x=2$	0.281189	56.2378
$x=3$	0.30162	60.324
$x=4$	0.181996	36.3992
$x=5$	0.058567	11.7134
$x=6$	0.00785299	1.57

x	0	1	2	3	4	5	6
$f(x)$	13	25	52	58	32	16	4
Exp $f(x)$	6	28	56	60	36	12	2

17/11/23

A random Variable X is said to be a ~~random~~ poission distribution if it assumes only on negitive values and its probability destribution is given by $P(x,\lambda) = P(X=\Omega) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{r}}{r!}, & r = 0,1,2,3 \ldots \infty \end{cases}$

Here, $\lambda > 0$ is

Conditions:

Can be derived as a limitting case of the binomical destribution Under the following conditions.

- The Probability of occurance of an event P is very small.
- No. of tries n is very large.
- np is a finite quanitity
  Say $\lambda = np$ then λ is parameter of the poission destribution.

Mean of the Poission Destribution:

$$P(X=r) = P(r) = \frac{e^{-\lambda}\lambda^{r}}{r!}, \quad r = 0,1,2,\ldots$$

$$\text{Mean } (\mu) = E(X) = \sum_{r=0}^{\infty} r\,P(r)$$

$$= \sum_{r=0}^{\infty} \frac{r e^{-\lambda}\lambda^{r}}{r!} = e^{-\lambda}\sum_{r=0}^{\infty} \frac{\lambda^{r}}{(r-1)!}$$

$$= e^{-\lambda} \sum_{r=0} \frac{\lambda^{r-1} \lambda}{(r-1)!}$$

$$= \lambda e^{-\lambda} \left[\sum_{r=0}^{\infty} \frac{\lambda^{r-1}}{(r-1)!} \right]$$

$$= \lambda e^{-\lambda} \left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots \right]$$

$$= \lambda e^{-\lambda} [e^{\lambda}]$$

$$\because e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$\mu = \lambda \frac{\cancel{e^{\lambda}}}{\cancel{e^{\lambda}}} = \lambda$$

$$\boxed{\mu = \lambda}$$

Variance of the Poission Destribution :

By definition, $P(X = r) = P(r) = \dfrac{e^{-\lambda} \lambda^r}{r!}, r = 0, 1, 2 \ldots$

$$\sigma^2 = \sum_{r=0}^{\infty} r^2 p(r) - \mu^2$$

$$\sigma^2 = \sum_{r=0}^{\infty} [r(r-1) + r] p(r) - \mu^2$$

$$= \sum_{r=0}^{\infty} r(r-1) \frac{e^{-\lambda} \lambda^r}{r!} + \sum_{r=0}^{\infty} r \frac{e^{-\lambda} \lambda^r}{r!} - \lambda^2$$

$$= e^{-\lambda} \sum_{r=0}^{\infty} \frac{\lambda^r}{(r-2)!} + \lambda - \lambda^2$$

$$= \lambda^2 e^{-\lambda} \left[\sum_{r=0}^{\infty} \frac{\lambda^{r-2}}{(r-2)!} \right] + \lambda - \lambda^2$$

$$= \lambda^2 e^{-\lambda} \left[0 + 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots \right] + \lambda - \lambda^2$$

$$\lambda^2 e^{-\lambda}[e^{\lambda}) + \lambda - \lambda^2$$

$$= \lambda^2 \frac{\cancel{e^{\lambda}}}{\cancel{e^{\lambda}}} + \lambda - \lambda^2$$

$$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2}$$

$$\sigma^2 = \lambda$$

$$\boxed{\sigma^2 = \lambda}$$

Q) If the probability that an individual suffers a bad reaction from a certain injection is 0.001. determine the probability that out of 2000 individuals

i) exactly 3   ii) > 2   iii) None.   iv) > 1

Given: $P = 0.001.$

$q = 0.999$

$n = 2000$

$\because \lambda = np = 2000 \times 0.001 = 2.$

$$\boxed{\lambda = 2}$$

Probability of bad reaction for exactly 3 individuals

$\Rightarrow P(X=3) =$

$$\because P(X=r) = P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$$

$$P(3) = \frac{e^{-2} 2^3}{3!}$$

$$= 0.1804$$

ii) $P(X>2) = 1 - [P(X=0) + P(X=1)] + P(X=2)]$

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

$$1 - \left[\frac{e^{-2} 2^0}{0!} + \frac{e^{-2} 2^1}{1!} + \frac{e^{-3} 2^2}{2!} \right]$$

$$= 1 - [0.13533 + 0.2706 + 0.27067]$$

$$= 1 - 0.6766$$

$P(X>2) = 0.323323$

iii) $P(X=0) = \frac{e^{-2} 2^0}{0!}$

$$= 0.13533$$

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

$$= 1 - 0.13533 - 0.2706$$

$$1 - 0.40600$$

~~$P(X>1) = 0.86466$~~

$P(X>1) = 0.59399$

$\therefore$ $P(X=3) = 0.1804$

$P(X>2) = 0.3233$

$P(X=0) = 0.1353$

$P(X>1) = 0.5939$

A Hospital Switch Board recieves an avarage of 4 emergency calls in 10 min interval. What is the probability that i) atmost 2 emergency calls in 10 mins. ii) $=3$ calls in 10 mins

$\lambda = 4$, calls / 10 mins

$P(X \leq 2) = P(0) + P(1) + P(2) = 0.2381$

$P(X=3) = 0.19536$

If a random Variable has poisson destribution such that P(1) = P(2). i) find Mean of the poisson Destribution. ii) P(4)

iii) $P(x \geq 1)$

iv $P(1 < x < 4)$

~~sign~~ Given: $P(1) = P(2)$

$\therefore P(X = r) = P(r) = \dfrac{e^{-\lambda}\lambda^{r}}{r!}$ (or)

By reccurence relation,

$\Rightarrow \dfrac{e^{-\lambda}\lambda^{1}}{1!} = \dfrac{e^{-\lambda}\lambda^{2}}{2!}$

$P(X+1) = \dfrac{\lambda}{x+1} \cdot P(x)$

$P(2) = \dfrac{\lambda}{2} P(1)$

$\dfrac{\lambda}{\lambda^2} = \dfrac{1}{2}$

$\dfrac{P(2)}{P(1)} = \dfrac{\lambda}{2}$

$\Rightarrow \lambda = 2$

$1 = \dfrac{\lambda}{2}$

$\lambda = 2$

$\therefore$ Mean of the poisson distribution $\lambda = 2$.

ii) $P(4) = \dfrac{e^{-2}\, 2^{4}}{3!} = 0.09022$

iii) $P(x \geq 1) = 1 - P(x < 1)$

$= 1 - P(x=0) = 1 - \dfrac{e^{-\lambda}2^{0}}{0!}$

$= 1 - 0.13533$

$= 0.80466$

iv) $P(1 < x < 4) = P(x=2) + P(X=3)$

$\dfrac{e^{-\lambda}2^{2}}{2!} + \dfrac{e^{-\lambda}2^{3}}{3!} = 0.27067 + 0.180447.$

$= 0.45111$

2% of the items of a factory are defective. The items are packed in boxes. What is the Probability that there will be (i) 2 defficitive ~~items~~ items (ii) atleast 3 ~~deff~~ defictive items in a box of 100 items.

$p = 0.02$ (defictive items $\frac{2}{100}$)

$q = \cancel{0.08}\ 0.98$

$n = 100$.

(i) $P(X=r) = P(r) = {}^{n}C_{r}\, p^{r} q^{n-r}$

~~ii) P(X = 2) =~~ (or)

with Poission Destribution

$\lambda = np = 2$

$$P(X=r) = P(r) = \frac{e^{-\lambda}\lambda^{r}}{r!}$$

$$P(X=2) = \frac{e^{-2}\,2^{2}}{2!} = 0.27067$$

$$P(x \geq 3) = 1 - P(X<3)$$

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

$$= 1 - \left[\cancel{0.135333}\ \frac{e^{-2}2^{0}}{0!} + \frac{e^{-2}2^{1}}{1!} + \frac{e^{-2}2^{2}}{2!}\right.$$

$$= 1 - [0.13533 + 0.27067 + 0.27$$

$$= 1 - 0.67667$$

$$= 0.32332$$

H.W If 2% of light bulbs are defective, find

i) one is defective (atleast) $1-(0.13533+0.27067)$ $1-0.40600 = 0.59399$

ii) exactly 7 are ~~to~~ defect. 0.003437

iii) Almost two are defective. 0.40600 ~~0.59399~~

iv) Probability of $(1 < x < 8)$ $P(1 < x < 8)$

in a sample of 100.

Using reccurence relation formula find the probability for $x = 0, 1, 2, 3, 4, 5$ if the mean of Poisson destribution is 3.

Given Mean of PD = 3

$$\boxed{\lambda = 3.}$$

By the definition of PD

$$P(x=r) = P(r) = \frac{e^{-\lambda}\lambda^{r}}{r!}.$$

$$P(r) = \frac{e^{-3}\,3^{r}}{r!}$$

$$P(x=0) = P(0) = \frac{e^{-3}\,3^{0}}{0!}$$

$$= 0.04978$$

By reccurence formula,

$$P(x+1) = \frac{\lambda}{x+1} P(x)$$

$$\because \lambda = 3$$

$$P(x+1) = \frac{3}{x+1} P(x)$$

$$P(x+1) = \frac{3}{x+1} P(0)$$

i) If $x=0$,

$\because P(0) = 0.0498$

$$P(1) = \frac{3}{1} P(0)$$

$P(1) = 0.1493612$

ii) If $x=1$.

$$P(2) = \frac{3}{1+1} P(1)$$

$= \frac{3}{2} P(1)$ $[\because P(1) = 0.149]$

$= 0.2240418$

iii) If $x=2$

$$P(3) = \frac{3}{2+1} (P(2))$$ $[\because P(2) = 0.2240]$

$= \frac{3}{3} P(2) = P(2)$

$= 0.2240418$

iv) $[\because P(3) = 0.2240]$

If $x=3$

$$P(4) = \frac{3}{3+1} P(3)$$

$= \frac{3}{4} P(3)$

$= 0.168031$

v)

If $x=4$

$$P(5) = \frac{3}{4+1} P(4)$$

$= \frac{3}{5} (P(4))$

$= 0.100818$

$\therefore$ $P(0) = 0.0498$
$P(1) = 0.1493612$
$P(2) = 0.2240418$
$P(3) = 0.2240418$
$P(4) = 0.168031$
$P(5) = 0.100818$.

18/11/23

If a poisson distribution is such that $P(x=1)\frac{3}{2} = P(x=3)$

find i) $P(x \geq 1)$ ii) $P(x \leq 3)$

iii) $P(2 \leq x \leq 5)$

$$\because P(x=r) = \frac{e^{-\lambda}\lambda^{r}}{r!}$$

$$P(x=1)\ \frac{3}{2} = P(x=3)$$

$$\frac{e^{-\lambda}\lambda^1}{1!}\ \frac{3}{2} = \frac{e^{-\lambda}\lambda^3}{3!}$$

$$\boxed{\lambda = 3}$$

i) $P(x \geq 1) = \cancel{P(x=0)+P(x=1)}\ 1 - P(x<1)$

$1 - P(x=0)$

$= 1 - \frac{e^{-\lambda}\lambda^0}{0!}$

$= 1 - 0.049787$

$= \cancel{1-0.950212}$

$= 0.950212.$

ii) $P(x \leq 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$

$= 0.04974 + 0.14936 + 0.2240 + 0.2240$

$= 0.64706.$

iii) $P(2 \leq x \leq 5) = P(2) + P(3) + P(4) + P(5)$

$= 0.22404 + 0.22404 + 0.168031 + 0.100818$

$= 0.716929.$

$\therefore P(x \geq 1) = 0.950212$

$P(x \leq 3) = 0.64706$

$P(2 \leq x \leq 5) = 0.716929.$

Fitting a ~~Bino~~ Poisson Destribution:

The destribution of typing mistakes commited by a person is given below. Assuming the destribution to be poission. Also find expected frequencies.

x	0	1	2	3	4	5
$f(x)$	42	33	14	6	4	1

Total frequencies $N = \sum f_i = 42+33+14+6+4+1$

$= 100$

$$\text{Mean} = \mu = \frac{\sum f_i x_i}{N} = \frac{1(33)+2(33)+14(2)+3(6)+4(4)+5(1)}{100}$$

$$= \frac{100}{100} = 1$$

For Fitting P.D;

$\mu = \lambda = 1$ for poission destribution.

Since, P.D of $x = P(x=r) = \dfrac{e^{-\lambda}\lambda^{r}}{r!}$

$$P(x=r) = \frac{e^{-1}\,1^{r}}{r!} \quad \because \lambda = 1.$$

$X = x$	$P(X=x)$	Exp frequency $NP(x)$
$X=0$	0.36787	36.787
$X=1$	0.36787	36.787
$X=2$	0.183939	18.3939
$X=3$	0.061313	6.1313
$X=4$	0.0153283	1.53283
$X=5$	[illegible]	[illegible]06566

Since, frequencies are always integers, by converting them to nearest integers, we get expected frequency.

Expected frequency :

37, 37, 18, 6, 2, 0.

Method of least square

Therefore,

x	0	1	2	3	4	5
observed $f(x)$	42	33	14	6	4	1
Expected $f(x)$	37	37	18	6	2	0

Fit a poisson destribution for the fowing and calulate expected frequency.

x	0	1	2	3	4
$f(x)$	109	65	22	3	1

$$N = \Sigma f_i = 200$$

$$M = \lambda = \frac{\Sigma f_i x}{\Sigma f_i} = \frac{122}{200} = 0.61$$

$X = x$	$P(X = x)$	Exp $f(x)$
$X = 0$	0.54335	108.67
$X = 1$	0.331444	66.288
$X = 2$	0.101090	20.218
$X = 3$	0.02055	4.11
$X = 4$	0.003134645	0.6264
		199.9129

x	0	1	2	3	4	
obtained $f(x)$	109	65	22	3	1	/200
Expected $f(x)$	109	66	20	4	1	/200