Problems on Joint Probability Mass function

Q). The joint probability mass function of (X, Y) is given by $p(x,y) = K(2x+3y)$, $x = 0,1,2$; $y = 1,2,3$. Find all the marginal and conditional distributions. Also find the probability distribution of $X+Y$, $2X-3Y$.

Soln: Given

Y \ X	0	1	2
1	3K	5K	7K
2	6K	8K	10K
3	9K	11K	13K

To find 'K'

$\because \sum\sum P(X=x, Y=y) = 1$,

$72K = 1 \Rightarrow \boxed{K = 1/72}$

Marginal distribution of X

X :	0	1	2
$P(X=x)$:	$18K = 18/72$	$24K = 24/72$	$30K = 30/72$

$\Rightarrow$

X :	0	1	2
$P(X=x)$:	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{5}{12}$

Marginal distribution of Y

Y :	1	2	3
$P(Y=y)$:	$15K = 15/72$	$24K = 24/72$	$33K = 33/72$

$\Rightarrow$

Y :	1	2	3
$P(Y=y)$:	$\frac{15}{72}$	$\frac{24}{72}$	$\frac{33}{72}$

Conditional distributions

Conditional distribution of x given $y=1$

x :	0	1	2
$P(X=x/Y=1)$:	$1/5$	$\frac{1}{3}$	$\frac{7}{15}$

$P(X=x/Y=1)$
$= \frac{P(X=x, Y=1)}{P(Y=1)}$

Conditional distribution of x given $y=2$

X :	0	1	2
$P(X=x/Y=2)$:	$1/4$	$1/3$	$5/12$

$P(X=x/Y=2)$
$= \frac{P(X=x, Y=2)}{P(Y=2)}$

Conditional distribution of x given $y=3$

X :	0	1	2
$P(X=x/Y=3)$:	$\frac{3}{11}$	$\frac{1}{3}$	$\frac{13}{33}$

$P(X=x/Y=3)$
$= \frac{P(X=x, Y=3)}{P(Y=3)}$

Conditional distribution of $\cancel{X}$ given $X=0$

Y :	1	2	3
$P(Y=y/X=0)$:	$\frac{1}{6}$	$\frac{1}{3}$	$1/2$

$P(Y=y/X=0)$
$= \frac{P(X=0, Y=y)}{P(X=0)}$

Conditional distribution of Y given $X=1$

Y :	1	2	3
$P(Y=y/X=1)$:	$5/24$	$8/24$	$11/24$

$P(Y=y/X=1)$
$= \frac{P(X=1, Y=y)}{P(X=1)}$

Conditional distribution of Y given $X=2$

Y :	1	2	3
$P(Y=y/X=2)$:	$7/30$	$10/30$	$13/30$.

Probability distribution of $X+Y$

overall
Possible pairs $= \{(0,1), (0,2), (0,3), (1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\}$.

Possible values (of $X+Y$) $= 1, 2, 3, 4, 5$.

$Z = X+Y$	Possibility	$P(Z = X+Y)$
1	$(0,1)$	$3K = 3/72$
2	$(0,2), (1,1)$	$6K + 5K = 11K = 11/72$
3	$(0,3), (1,2), (2,1)$	$9K + 8K + 7K = 24K = 24/72$
4	$(1,3), (2,2)$	$11K + 10K = 21K = \frac{21}{72}$
5	$(2,3)$	$13K = 13/72$.

$\therefore$ Prob. distribution of $Z = X+Y$

$Z = z$	1	2	3	4	5
$P(Z=z)$	$3/72$	$11/72$	$24/72$	$21/72$	$13/72$

Probability distribution of $2X - 3Y$

Possible values (of $2X-3Y$) $= -3, -6, -9, -1, -4, -7, 1, -2, -5$.

$Z = 2X - 3Y$	Possibility	$P(Z = 2X - 3Y)$
-3	$(0,1)$	$3/72$
-6	$(0,2)$	$6/72$
-9	$(0,3)$	$9/72$
-1	$(1,1)$	$5/72$

-4	$(1, 2)$	$8/72$
-7	$(1, 3)$	$11/72$
1	$(2, 1)$	$7/72$
-2	$(2, 2)$	$10/72$
-5	$(2, 3)$	$13/72$

$\therefore$ Prob. distribution of $Z = 2X - 3Y$

$Z = z$:	-3	-6	-9	-1	-4	-7	1	-2	-5
$P(Z = z)$:	$\frac{3}{72}$	$\frac{6}{72}$	$\frac{9}{72}$	$\frac{5}{72}$	$\frac{8}{72}$	$\frac{11}{72}$	$\frac{7}{72}$	$\frac{10}{72}$	$\frac{13}{72}$

---

02) The joint probability distribution of a two dimensional discrete random variable (X, Y) is

Y \ X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

Find (a) $P(X \le 2)$ (b) $P(Y > 1)$ (c) $P(X \le 1, Y \le 1)$

(d) $P(X > 3 / Y = 2)$ (e) $P(X > 4 / Y < 2)$ (f) $P(X > Y)$.

(g) $P\{\max(X, Y) = 3\}$ (h) $P(X + Y \le 3)$

(i) the probability distribution of $Z = \min(X, Y)$

Marginal distribution of X

x:	0	1	2	3	4	5
$P(X=x)$:	0.03	0.08	0.16	0.21	0.24	0.28

Marginal distribution of Y

y:	0	1	2	3
$P(Y=y)$:	0.25	0.26	0.25	0.24

(a) $P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$= 0.03 + 0.08 + 0.16 = 0.27$.

$\therefore \boxed{P(X \le 2) = 0.27}$

(b) $P(Y > 1) = P(Y=2) + P(Y=3) = 0.25 + 0.24 = 0.49$

$\therefore \boxed{P(Y > 1) = 0.49}$

(c) $P(X \le 1, Y \le 1) = P(X = 0, 1; Y = 0, 1)$

$= P(0,0) + P(0,1) + P(1,0) + P(1,1)$

$= 0 + 0.01 + 0.01 + 0.02 = 0.04$

$\therefore \boxed{P(X \le 1, Y \le 1) = 0.04}$

(d) $P(X > 3 / Y = 2) = \dfrac{P(X > 3, Y = 2)}{P(Y = 2)}$ —— (1)

$P(X > 3, Y = 2) = P(X = 4, 5; Y = 2) = P(4,2) + P(5,2)$

$= 0.05 + 0.06 = 0.11$

$P(Y = 2) = 0.25$

$\boxed{P(X > 3 / Y = 2) = \dfrac{0.11}{0.25} = 0.44}$

3

(e) $P(X > 4 \mid Y < 2) = \dfrac{P(X > 4, Y < 2)}{P(Y < 2)}$ —— (1)

$P(X > 4, Y < 2) = P(X = 5, Y = 0, 1)$

$= P(5, 0) + P(5, 1) = 0.09 + 0.08 = 0.17.$

$P(Y < 2) = P(Y = 0) + P(Y = 1) = 0.25 + 0.26 = 0.51$

$\therefore \boxed{P(X > 4 \mid Y < 2) = \dfrac{0.17}{0.51} = 0.33}$

(f) To find $P(X > Y)$

Overall pairs : $\{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (1,3), (2,0), (2,1), (2,2), (2,3), (3,0), (3,1), (3,2), (3,3), (4,0), (4,1), (4,2), (4,3), (5,0), (5,1), (5,2), (5,3)\}$

Favourable pairs $(X > Y)$: $\{(1,0), (2,0), (2,1), (3,0), (3,1), (3,2), (4,0), (4,1), (4,2), (4,3), (5,0), (5,1), (5,2), (5,3)\}$

$\therefore P(X > Y) = 0.01 + 0.03 + 0.04 + 0.05 + 0.05 + 0.05 + 0.07 + 0.06 + 0.05 + 0.06 + 0.09 + 0.08 + 0.0$
$+ 0.05$
$= 0.75$

$\therefore \boxed{P(X > Y) = 0.75}$

(g) To find $P\{\max(x, y) = 3\}$

Favourable pairs $\max(x,y) = 3$ $= \{(0,3), (1,3), (2,3), (3,0), (3,1), (3,2), (3,3)\}$

$\therefore P\{\max(x,y) = 3\} = 0.01 + 0.02 + 0.04 + 0.05 + 0.05 + 0.05 + 0.06.$

$= 0.28$

(h) To find $P(x+y \leq 3)$

Favourable pair $x+y \leq 3$ $= \{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (2,0), (2,1), (3,0)\}$

$\therefore P(x+y \leq 3) = 0 + 0.01 + 0.01 + 0.01 + 0.01 + 0.02 + 0.03 + 0.03 + 0.04 + 0.05$

$= 0.21.$

(i) Probability distribution of $z = \min(x, y)$

Possible values : 0, 1, 2, 3.

$z = \min(x,y)$	Possibilities	Probability
0	(0,0), (0,1), (0,2), (0,3), (1,0), (2,0), (3,0), (4,0), (5,0)	0.28
1	(1,1), (1,2), (1,3), (2,1), (3,1), (4,1), (5,1)	0.30
2	(2,2), (2,3), (3,2), (4,2), (5,2)	0.25
3	(3,3), (4,3), (5,3)	0.17

Probability distribution of $Z = \min(X, Y)$

Z:	0	1	2	3
$P(Z=z)$:	0.28	0.30	0.25	0.17

---

03) Let X and Y be integer valued R.V's with

$P(X=m, Y=n) = q^2 p^{m+n-2}$, $m, n = 1, 2, 3, \ldots$,

$p + q = 1$. Are X and Y independent.

Soln: for convenience let $m = x$, $n = y$

$$\therefore P(X=x, Y=y) = q^2 p^{x+y-2}, \quad x, y = 1, 2, 3, \ldots$$

$$P(X=x) = \sum_{y=1}^{\infty} P(X=x, Y=y)$$

$$= \sum_{y=1}^{\infty} q^2 p^{x+y-2} = q^2 p^{x-2} \sum_{y=1}^{\infty} p^y$$

$$= q^2 p^{x-2} \{p + p^2 + p^3 + \ldots\}$$

$$= q^2 p^{x-2} \cdot p[1 + p + p^2 + \ldots]$$

$$= q^2 p^{x-1} \cdot (1-p)^{-1} = q^2 p^{x-1} \cdot q^{-1}$$

$$P(X=x) = q p^{x-1}, \quad x = 1, 2, 3, \ldots$$

|||ly $P(Y=y) = q p^{y-1}$, $y = 1, 2, 3, \ldots$

Now $P(X=x) \times P(Y=y) = q p^{x-1} \cdot q p^{y-2}$

$$= q^2 p^{x+y-2}$$

$$= P(X=x, Y=y).$$

$\therefore$ X and Y are independent.

Sums for Practice

01) Given the joint probability mass function below

Y \ X	1	2	3	4	5	6
2	3K	5K	K	K	2K	0
4	0	2K	3K	5K	K	K
6	K	2K	K	2K	K	2K

(a) Find the marginal distributions

(b) Find the conditional distributions of X given Y = 2, Y given X = 4.

(c) $P(X \leq 3, Y \leq 4)$ (d) $P(X \leq 2)$ (e) $P(Y \leq 4)$

(f) $P(X \leq 1 / Y \leq 4)$ (g) $P(X+Y \leq 4)$ (h) $P(X+Y > 10)$

(i) Probability distribution of $Z = \max(X, Y)$

Problems based on Joint probability density function

Note:

Data 1: The limits of 'x' & 'y' are independent

(ie) (a) $x \geq 0,\ y \geq 0$ (b) $0 \leq x, y < \infty$ (ie $0 \leq x < \infty,\ 0 \leq y < \infty$)

Data 2: The limits of 'x' & 'y' are dependent.

(ie) (a) $0 \leq x < y \leq 1$ (b) $|x| < y$

---

01) The joint pdf of a two dimensional random variable (X, Y) is given by $f(x, y) = xy^2 + \frac{x^2}{8}$, $0 \leq x \leq 2$, $0 \leq y \leq 1$.

Find (i) $P(Y < \frac{1}{2})$ (ii) $P(X > 1)$ (iii) $P(X > 1 / Y < \frac{1}{2})$

(iv) $P(Y < \frac{1}{2} / X > 1)$ (v) $P(Y < \frac{1}{2} / X = 1)$

(vi) $P(X < Y)$ (vii) $P(X + Y \leq 1)$ (viii) $P(X > Y)$

(ix) $P(X + Y > 1)$.

Soln:

Given $f(x, y) = xy^2 + \frac{x^2}{8}$, $0 \leq x \leq 2$, $0 \leq y \leq 1$.

Marginal density function of X

$$f(x) = \int_0^1 f(x, y)\, dy = \int_0^1 \left(xy^2 + \frac{x^2}{8}\right) dy$$

$$= \left(\frac{xy^3}{3} + \frac{x^2 y}{8}\right)_0^1 = \left(\frac{x}{3} + \frac{x^2}{8}\right) - (0)$$

$$\therefore \boxed{f(x) = \frac{x}{3} + \frac{x^2}{8},\quad 0 \leq x \leq 2}$$

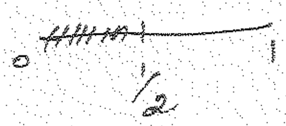
Marginal density function of Y

$$f(y) = \int_0^2 f(x,y)\,dx = \int_0^2 \left(xy^2 + \frac{x^2}{8}\right) dx$$

$$= \left(\frac{x^2y^2}{2} + \frac{x^3}{24}\right)_0^2 = \left(\frac{4y^2}{2} + \frac{8}{24}\right) - (0) = 2y^2 + \tfrac{1}{3}.$$

$$\therefore \boxed{f(y) = 2y^2 + \frac{1}{3}, \quad 0 \le y \le 1}$$

ⓐ $$P(Y < \tfrac{1}{2}) = \int_0^{1/2} f(y)\,dy$$

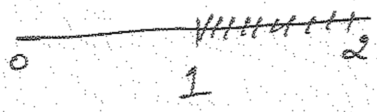


$$= \int_0^{1/2} \left(2y^2 + \frac{1}{3}\right) dy = \left(\frac{2y^3}{3} + \frac{y}{3}\right)_0^{1/2}$$

$$= \left[\frac{2}{3}\left(\frac{1}{8}\right) + \frac{1}{3}\left(\frac{1}{2}\right)\right] - (0) = \frac{1}{12} + \frac{1}{6} = \frac{1}{4}.$$

$$\therefore \boxed{P(Y < \tfrac{1}{2}) = \frac{1}{4}}$$

ⓑ $$P(X > 1) = \int_1^2 f(x)\,dx$$



$$= \int_1^2 \left(\frac{x}{3} + \frac{x^2}{8}\right) dx = \left(\frac{x^2}{6} + \frac{x^3}{24}\right)_1^2$$

$$= \left(\frac{4}{6} + \frac{8}{24}\right) - \left(\frac{1}{6} + \frac{1}{24}\right) = \frac{19}{24}.$$

$$\therefore \boxed{P(X > 1) = \frac{19}{24}}$$

(e) $P(Y < \frac{1}{2} \mid X = 1) = \left\{ \int_0^{1/2} f(y/x)\, dy \right\}_{x=1}$ —— (1)

Here $f(y/x) = \dfrac{f(x,y)}{f(x)} = \dfrac{xy^2 + x^2/8}{x/3 + x^2/8}$

$\therefore \int_0^{1/2} f(y/x)\, dy = \int_0^{1/2} \dfrac{xy^2 + x^2/8}{x/3 + x^2/8}\, dy$

$= \dfrac{1}{x/3 + x^2/8} \left\{ \dfrac{xy^3}{3} + \dfrac{x^2 y}{8} \right]_0^{1/2}$

$= \dfrac{1}{\frac{x}{3} + \frac{x^2}{8}} \left\{ \dfrac{x}{3}\left(\dfrac{1}{8}\right) + \dfrac{x^2}{8}\left(\dfrac{1}{2}\right) \right]$

$= \dfrac{1}{x/3 + x^2/8} \left\{ \dfrac{x}{24} + \dfrac{x^2}{16} \right]$.

$\therefore$ (1) $\Rightarrow P(Y < \frac{1}{2} \mid X = 1) = \dfrac{1}{(1/3 + 1/8)} \left(\frac{1}{24} + \frac{1}{16} \right)$

$= \dfrac{1}{11/24} \left(\dfrac{5}{48} \right) = 5/22.$

$\therefore \boxed{P(Y < \frac{1}{2} \mid X = 1) = \frac{5}{22}}$

(c) $P(X>1 / Y<\frac{1}{2}) = \dfrac{P(X>1, Y<\frac{1}{2})}{P(Y<\frac{1}{2})}$ —— (1)

Now $P(X>1, Y<\frac{1}{2}) = \int_0^{1/2}\int_1^2 f(x,y)\,dx\,dy$

$= \int_0^{1/2}\int_1^2 \left(xy^2 + \frac{x^2}{8}\right) dx\,dy$ —— (2).

Now $\int_1^2 \left(xy^2 + \frac{x^2}{8}\right) dx = \left(\frac{x^2y^2}{2} + \frac{x^3}{24}\right)_1^2$

$= \left(\frac{4y^2}{2} + \frac{8}{24}\right) - \left(\frac{y^2}{2} + \frac{1}{24}\right) = \frac{3y^2}{2} + \frac{7}{24}$

(2) $\Rightarrow P(X>1, Y<\frac{1}{2}) = \int_0^{1/2}\left(\frac{3y^2}{2} + \frac{7}{24}\right) dy$

$= \left(\frac{\cancel{3}}{2}\left(\frac{y^3}{\cancel{3}}\right) + \frac{7y}{24}\right)_0^{1/2} = \left(\frac{y^3}{2} + \frac{7y}{24}\right)_0^{1/2}$

$= \frac{1}{2}\left(\frac{1}{8}\right) + \frac{7}{24}\left(\frac{1}{2}\right) = \frac{1}{16} + \frac{7}{48} = \frac{5}{24}$.

$\therefore P(X>1, Y<\frac{1}{2}) = 5/24$

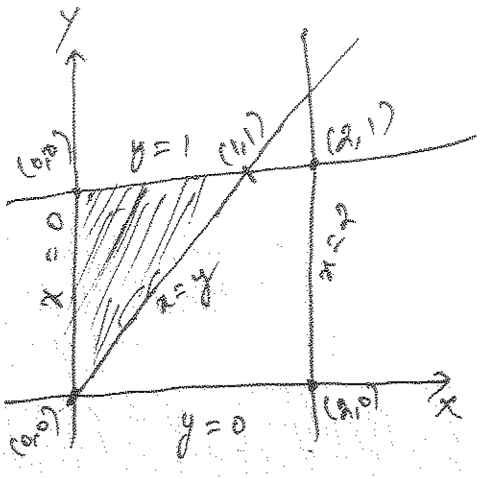
Also $P(Y<\frac{1}{2}) = \frac{1}{4}$ (from (a))

$\therefore P(X>1 / Y<\frac{1}{2}) = \dfrac{5/24}{1/4} = \frac{5}{6}$.

(d) $P(Y<\frac{1}{2} / X>1) = \dfrac{P(X>1, Y<\frac{1}{2})}{P(X>1)} = \dfrac{5/24}{19/24} = \frac{5}{19}$

$\therefore \boxed{P(Y<\frac{1}{2} | X>1) = \frac{5}{19}}$

(f) $P(X<Y) = \int_0^1 \int_0^y f(x,y)\,dx\,dy$



$= \int_0^1 \int_0^y \left(xy^2 + \frac{x^2}{8}\right) dx\,dy$ — (1)

Now, $\int_0^y \left(xy^2 + \frac{x^2}{8}\right) dx = \left(\frac{x^2y^2}{2} + \frac{x^3}{24}\right)_0^y$

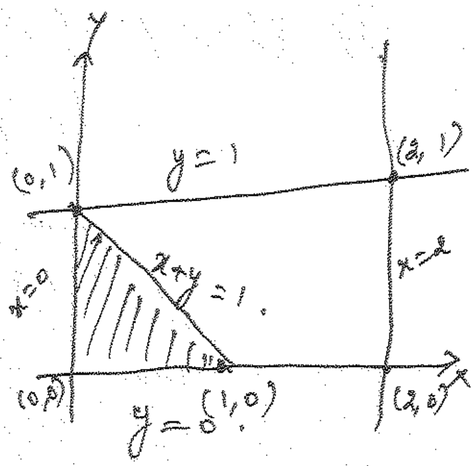
$= \frac{y^4}{2} + \frac{y^3}{24} - 0 = \frac{y^4}{2} + \frac{y^3}{24}$.

(1) $\Rightarrow$ $P(X<Y) = \int_0^1 \left(\frac{y^4}{2} + \frac{y^3}{24}\right) dy = \left(\frac{y^5}{10} + \frac{y^4}{96}\right)_0^1$

$= \frac{1}{10} + \frac{1}{96} = \frac{53}{480}$

$\therefore \boxed{P(X<Y) = \frac{53}{480}}$

(g) $P(X+Y \le 1) = \int_0^1 \int_0^{1-y} f(x,y)\,dx\,dy$



$= \int_0^1 \int_0^{1-y} \left(xy^2 + \frac{x^2}{8}\right) dx\,dy$ — (1)

$x+y=1$

x	0	1
y	1	0

Now $\int_0^{1-y} \left(xy^2 + \frac{x^2}{8}\right) dx = \left(\frac{x^2y^2}{2} + \frac{x^3}{24}\right)_0^{1-y}$

$= \frac{y^2}{2}(1-y)^2 + \frac{1}{24}(1-y)^3$.

$= \frac{y^2}{2}(1+y^2-2y) + \frac{1}{24}(1-3y+3y^2-y^3)$.

$$= \frac{y^2}{2} + \frac{y^4}{2} - y^3 + \frac{1}{24} - \frac{y}{8} + \frac{y^2}{8} - \frac{y^3}{24}.$$

$$= \frac{1}{24} - \frac{y}{8} + \frac{5y^2}{8} - \frac{25y^3}{24} + \frac{y^4}{2}.$$

(3) $\Rightarrow P(X+Y \le 1) = \int_0^1 \left(\frac{1}{24} - \frac{y}{8} + \frac{5y^2}{8} - \frac{25y^3}{24} + \frac{y^4}{2} \right) dy$

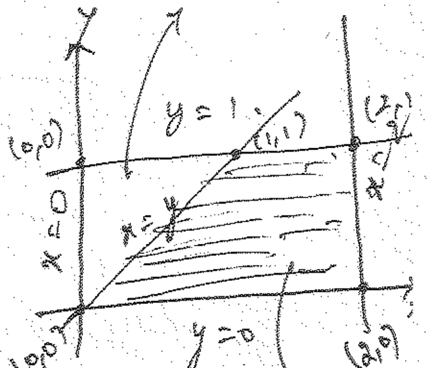
$$= \left(\frac{y}{24} - \frac{y^2}{16} + \frac{5y^3}{24} - \frac{25y^4}{96} + \frac{y^5}{10} \right)_0^1$$

$$= \frac{1}{24} - \frac{1}{16} + \frac{5}{24} - \frac{25}{96} + \frac{1}{10}$$

$$= \frac{13}{480}.$$

$\therefore \boxed{P(X+Y \le 1) = \frac{13}{480}}$

triangular,



Not a triangular surface. (Look for a triangular surface)

(h) $P(X > Y) = 1 - P(X \le Y)$

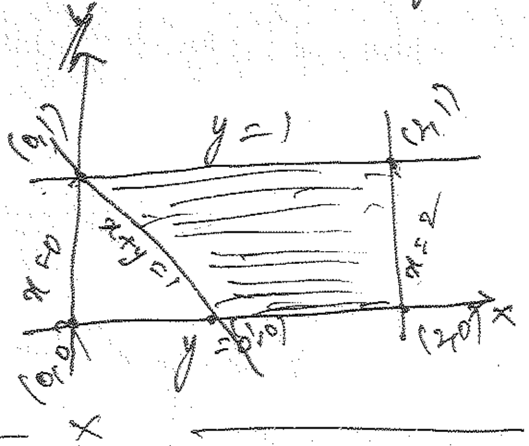
$$= 1 - \frac{53}{480} = \frac{427}{480}$$

$\therefore \boxed{P(X > Y) = \frac{427}{480}}$

(i) $P(X+Y \ge 1) = 1 - P(X+Y < 1)$

$$= 1 - \frac{13}{480} = \frac{467}{480}$$

$\therefore \boxed{P(X+Y \ge 1) = \frac{467}{480}}$



② If the joint distribution function of X and Y is given by $F(x,y) = (1-e^{-x})(1-e^{-y})$ for $x>0, y>0$.

(i) Find the marginal densities of X and Y.

(ii) Are X and Y independent.

(iii) Find $P(1<x<3, 1<y<2)$.

(iv) conditional densities.

Soln:

To find joint density function $f(x,y)$

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y) = \frac{\partial^2}{\partial x \partial y}(1-e^{-x})(1-e^{-y})$$

$$= \frac{\partial}{\partial x}(1-e^{-x})(e^{-y})$$

$$= e^{-x} e^{-y}.$$

$$\therefore \boxed{f(x,y) = e^{-x}.e^{-y}, \quad x>0, \ y>0}$$

ⓐ Marginal densities of X and Y

$$f(x) = \int_0^\infty f(x,y)\,dy = \int_0^\infty e^{-x}.e^{-y}\,dy$$

$$= e^{-x}(-e^{-y})_0^\infty = e^{-x}(0+1) = e^{-x}$$

$$\therefore \boxed{f(x) = e^{-x}, \quad x>0}$$

$$f(y) = \int_0^\infty f(x,y)\,dx = \int_0^\infty e^{-x}.e^{-y}\,dx$$

$$= e^{-y}(-e^{-x})_0^\infty = e^{-y}(0+1) = e^{-y}$$

$$\therefore \boxed{f(y) = e^{-y}, \quad y>0}$$

(b) X and Y are independent if $f(x) \cdot f(y) = f(x, y)$

Now $f(x) \cdot f(y) = e^{-x} \cdot e^{-y} = f(x, y)$

$\therefore$ X and Y are independent.

(c) $P(1 < X < 3, 1 < Y < 2) = \int_1^2 \int_1^3 f(x, y)\, dx\, dy$

$$= \int_1^2 \int_1^3 e^{-x} e^{-y}\, dx\, dy \quad \text{—— (1)}$$

Now $\int_1^3 e^{-x} \cdot e^{-y}\, dx = e^{-y}\left(-e^{-x}\right)_1^3$

$$= e^{-y}\left(-e^{-3} + e^{-1}\right)$$

(1) $\Rightarrow$ $P(1 < X < 3, 1 < Y < 2) = \left(e^{-1} - e^{-3}\right) \int_1^2 e^{-y}\, dy$

$$= \left(e^{-1} - e^{-3}\right)\left(-e^{-y}\right)_1^2 = \left(e^{-1} - e^{-3}\right)\left(-e^{-2} + e^{-1}\right)$$

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

(d) Conditional density function of X given Y

$$f(x/y) = \frac{f(x, y)}{f(y)} = \frac{e^{-x} \cdot e^{-y}}{e^{-y}} = e^{-x}$$

$$\therefore \boxed{f(x/y) = e^{-x}, \quad x > 0}$$

Conditional density function of Y given X

$$f(y/x) = \frac{f(x, y)}{f(x)} = \frac{e^{-x} \cdot e^{-y}}{e^{-x}} = e^{-y}$$

$$\therefore \boxed{f(y/x) = e^{-y}, \quad y > 0}$$

(3) Examine whether the variables X and Y are independent, whose joint density is

$f(x,y) = x e^{-x(y+1)}$, $0 < x, y < \infty$.

Soln:

Given $f(x,y) = x e^{-x(y+1)}$, $0 < x < \infty$, $0 < y < \infty$.

Now $f(x) = \int_0^\infty f(x,y)\,dy = \int_0^\infty x e^{-x(y+1)}\,dy$

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

$= (0) - (-e^{-x}) = e^{-x}$.

$\therefore \boxed{f(x) = e^{-x}, \quad 0 < x < \infty}$

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

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

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

$\therefore \boxed{f(y) = \frac{1}{(y+1)^2}, \quad 0 < y < \infty}$

Now $f(x) \cdot f(y) = \frac{e^{-x}}{(y+1)^2} \neq f(x,y)$

$\therefore$ X and Y are not independent

10

(4) The joint pdf of a R.V X and Y is given by $f(x,y) = kxy\, e^{-(x^2+y^2)}$, $x>0,\ y>0$. Find the value of k and prove also that X and Y are independent

Soln: To find 'k'

WKT $\int_0^\infty \int_0^\infty f(x,y)\, dx\, dy = 1$

$$\Rightarrow \int_0^\infty \int_0^\infty k\,xy\, e^{-x^2} \cdot e^{-y^2} \cdot dx\, dy = 1. \quad \text{——— (1)}$$

Now $\int_0^\infty k\,xy\, e^{-x^2} \cdot e^{-y^2} dx = ky\,e^{-y^2} \int_0^\infty x\, e^{-x^2} dx$

put $x^2 = t$
$2x\,dx = dt$
$x\,dx = \frac{dt}{2}$

$$= ky\,e^{-y^2} \int_0^\infty e^{-t} \frac{dt}{2}$$

$$= \frac{ky\,e^{-y^2}}{2} \left(-e^{-t}\right)_0^\infty$$

$$= \frac{ky\,e^{-y^2}}{2} (0+1) = \frac{ky\,e^{-y^2}}{2}$$

put $y^2 = t$
$2y\,dy = dt$
$y\,dy = \frac{dt}{2}$.

$$(1) \Rightarrow \int_0^\infty \frac{ky\,e^{-y^2}}{2}\, dy = 1$$

$$\Rightarrow \frac{k}{2} \int_0^\infty e^{-t} \frac{dt}{2} = 1.$$

$$\Rightarrow \frac{k}{4} \left(-e^{-t}\right)_0^\infty = 1$$

$$\Rightarrow k/4\,(0+1) = 1 \quad \Rightarrow k/4 = 1 \quad \Rightarrow \boxed{k = 4}$$

$$\therefore f(x,y) = 4xy\, e^{-(x^2+y^2)}, \quad x>0,\ y>0.$$
$$= 4xy\, e^{-x^2} \cdot e^{-y^2}$$

To prove x and y are independent

$$f(x) = \int_0^\infty f(x,y)\,dy = \int_0^\infty 4xy\, e^{-x^2} \cdot e^{-y^2} \cdot dy$$

put $y^2 = t$
$2y\,dy = dt$
$y\,dy = \frac{dt}{2}$

$$= 4x e^{-x^2} \int_0^\infty y e^{-y^2} dy$$
$$= \cancel{4}^{2} x e^{-x^2} \int_0^\infty e^{-t} \cdot \frac{dt}{\cancel{2}}$$
$$= 2x e^{-x^2} \left(-e^{-t}\right)_0^\infty = 2xe^{-x^2}(0+1)$$
$$\therefore \boxed{f(x) = 2xe^{-x^2}, \quad x>0}$$

$$f(y) = \int_0^\infty f(x,y)\,dx = \int_0^\infty 4xy\, e^{-x^2} \cdot e^{-y^2} \cdot dy$$

put $x^2 = t$
$2x\,dx = dt$
$x\,dx = \frac{dt}{2}$

$$= 4y e^{-y^2} \int_0^\infty x e^{-x^2} dx$$
$$= \cancel{4}^{2} y e^{-y^2} \int_0^\infty e^{-t} \frac{dt}{\cancel{2}}$$
$$= 2ye^{-y^2} \left(-e^{-t}\right)_0^\infty = 2ye^{-y^2}(0+1)$$
$$\therefore \boxed{f(y) = 2ye^{-y^2}, \quad y>0}$$

$$\therefore f(x) \cdot f(y) = \left(2xe^{-x^2}\right)\left(2ye^{-y^2}\right) = 4xy\,e^{-x^2} \cdot e^{-y^2}$$
$$= f(x,y)$$

$\therefore$ x and y are independent

Sums for Practice

① If the joint p.d.f of a random variable (X, Y) is given by $f(x, y) = x^2 + \frac{xy}{3}$, $0 \le x \le 1$, $0 \le y \le 2$.

ⓐ Find the conditional densities of X on Y & Y on X.

ⓑ Find $P(X < \frac{1}{2} \mid Y > 1)$. ⓒ $P(Y > 1 \mid X < \frac{1}{2})$

ⓓ $P(X < \frac{1}{2} \mid Y = 1)$. ⓔ $P(X < Y)$ ⓕ $P(X + Y > 1)$.

② The joint pdf of a bivariate R.V (X, Y) is given by

$f(x, y) = Kxy$, $0 < x, y < 1$.

ⓐ Find K, ⓑ $P(X + Y < 1)$ ⓒ Are X and Y independen

③ The joint probability distribution of X and Y is given by $f(x, y) = \frac{1}{8}(6 - x - y)$, $0 < x < 2$, $2 < y < 4$.

Find $f(y/x = 2)$, $P(X < 1 \mid Y < 3)$

④ The joint density function of two random variables X and Y is $f(x, y) = e^{-(x+y)}$, $0 \le x, y < \infty$.

Are X and Y independent. Find $P(X > Y)$, $P(X + Y < 1)$, $P(X < 1 \mid Y = 2)$.

⑤ If X and Y have joint p.d.f $f(x, y) = x + y$, $0 < x, y < 1$, check whether X and Y are independent.

12

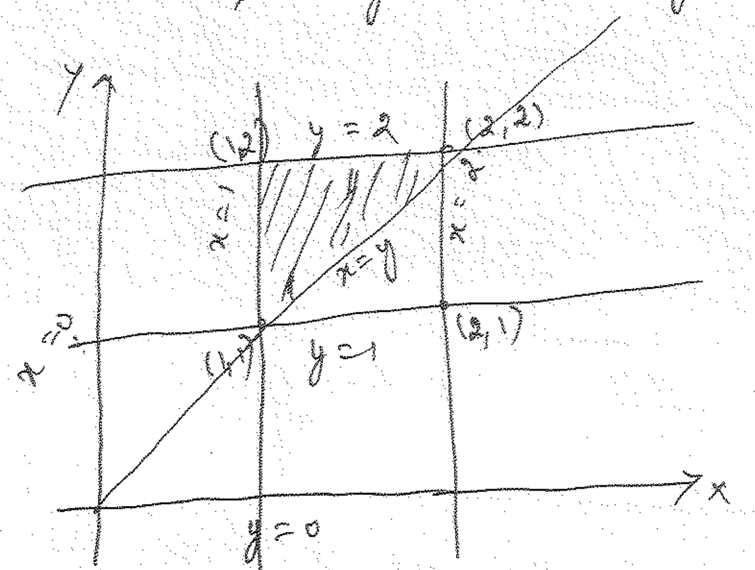
01) The joint probability density function of the two dimensional random variable is $f(x, y) = \frac{8xy}{9}$, $1 < x < y < 2$.

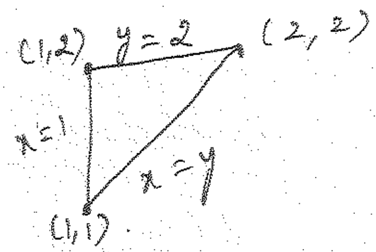
Find the conditional density function of X given Y and Y given X. Also verify the conditional density functions are valid. Find $P(X < 3/2)$.

Soln:

Given region of integration for $f(x, y)$:

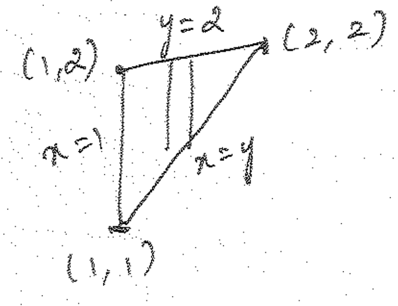
$1 < x < 2, \quad 1 < y < 2, \quad x < y$





∴ the region of integration is (for $f(x,y)$)

Marginal densities of X and Y

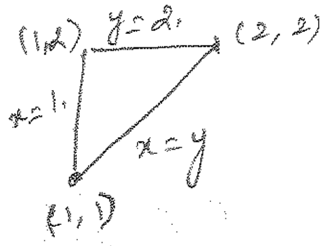


$$f(x) = \int_x^2 f(x, y)\, dy = \int_x^2 \frac{8xy}{9}\, dy$$

$$= \frac{\cancel{8}^{4}x}{9}\left(\frac{y^2}{\cancel{2}}\right)_x^2 = \frac{4x}{9}(4 - x^2)$$

$$\therefore \boxed{f(x) = \frac{4x}{9}(4 - x^2), \quad 1 \le x \le 2}$$

$$f(y) = \int_1^y f(x,y)\,dx = \int_1^y \frac{8xy}{9}\,dx$$



$$= \frac{\cancel{8}^{4}y}{9}\left(\frac{x^2}{\cancel{2}}\right)_1^y = \frac{4y}{9}(y^2-1).$$

$$\therefore \boxed{f(y) = \frac{4y}{9}(y^2-1),\quad 1 \le y \le 2}$$

Conditional density function of X given Y

$$f(x/y) = \frac{f(x,y)}{f(y)} = \frac{8xy/9}{\frac{4y}{9}(y^2-1)} = \frac{\cancel{8}^{2}x\cancel{y}}{\cancel{9}} \times \frac{\cancel{9}}{\cancel{4}\cancel{y}(y^2-1)}$$

$$\therefore \boxed{f(x/y) = \frac{2x}{y^2-1},\quad 1 < x < y}$$

Conditional density function of Y given X

$$f(y/x) = \frac{f(x,y)}{f(x)} = \frac{\frac{8xy}{9}}{\frac{4x}{9}(4-x^2)} = \frac{\cancel{8}^{2}\cancel{x}y}{\cancel{9}} \times \frac{\cancel{9}}{\cancel{4}\cancel{x}(4-x^2)}$$

$$\therefore \boxed{f(y/x) = \frac{2y}{4-x^2},\quad x < y < 2}$$

Verification

$$\int_1^y f(x/y)\,dx = \int_1^y \frac{2x}{y^2-1}\,dx = \frac{1}{y^2-1}\left(x^2\right)_1^y$$

$$= \frac{1}{y^2-1}(y^2-1) = 1$$

$\therefore f(x/y)$ is valid.

Also $\int_x^2 f(y/x)\, dy = \int_x^2 \frac{2y}{4-x^2}\, dx = \frac{1}{4-x^2}\left(y^2\right)_x^2$

$$= \frac{1}{4-x^2}(4-x^2) = 1$$

$\therefore f(y/x)$ is valid

$$P(X < 3/2) = \int_1^{3/2} f(x)\, dx = \int_1^{3/2} \frac{4x}{9}(4-x^2)\, dx$$

$$= \frac{4}{9}\int_1^{3/2}(4x - x^3)\, dx = \frac{4}{9}\left(\frac{\overset{2}{\cancel{4}}x^2}{\cancel{2}} - \frac{x^4}{4}\right)_1^{3/2}$$

$$= \frac{4}{9}\left\{\left[2\left(\frac{9}{4}\right) - \frac{1}{4}\left(\frac{81}{16}\right)\right] - \left[2 - \frac{1}{4}\right]\right\}$$

$$= \frac{4}{9}\left\{\frac{9}{2} - \frac{81}{64} - 2 + \frac{1}{4}\right\} = \frac{95}{144}$$

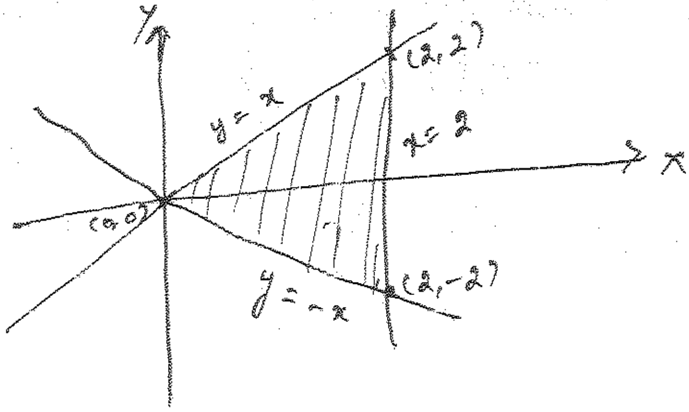
$$\therefore \boxed{P(X < 3/2) = \frac{95}{144}}$$

Q2) Given $f(x, y) = cx(x-y)$, $0 < x < 2$, $|y| < x$ and find

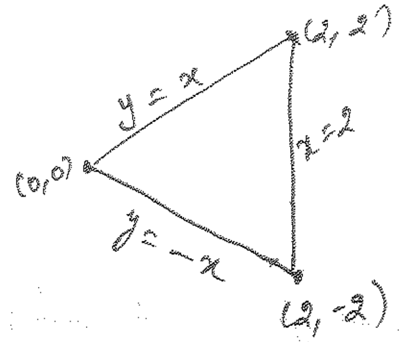
(i) c (ii) $f(x)$ & $f(y)$ (iii) $f(x/y)$ & $f(y/x)$

Soln: Given $f(x, y) = cx(x-y)$, $0 < x < 2$, $-x < y < x$

Region for $f(x, y)$



(a) To find 'c'



$$\int_0^2 \int_{-x}^{x} f(x, y)\, dy\, dx = 1$$

$$\Rightarrow \int_0^2 \int_{-x}^{x} c\,x(x-y)\, dy\, dx = 1 \quad \text{— (1)}$$

$$\text{Now } c\int_{-x}^{x} (x^2 - xy)\, dy = c\left[x^2 y - \frac{xy^2}{2}\right]_{-x}^{x}$$

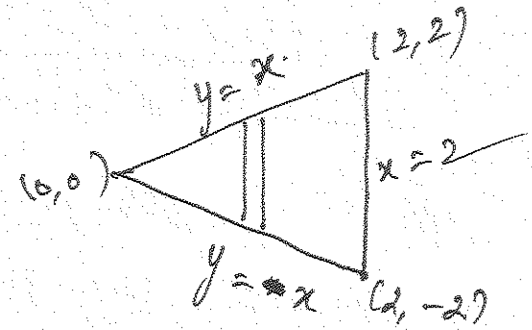
$$= c\left\{\left(x^3 - \frac{x^3}{2}\right) - \left(-x^3 - \frac{x^3}{2}\right)\right] =$$

$$= c\left(x^3 - \frac{x^3}{2} + x^3 + x^3/2\right) = 2cx^3$$

$$(1) \Rightarrow \int_0^2 2cx^3 = 1 \quad \Rightarrow 2c\left(\frac{x^4}{4}\right)_0^2 = 1$$

$$\Rightarrow 2c\left(\frac{16}{4} - 0\right) = 1 \quad \Rightarrow 8c = 1 \quad \Rightarrow \boxed{c = 1/8}$$

(b) To find f(x) & f(y)



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

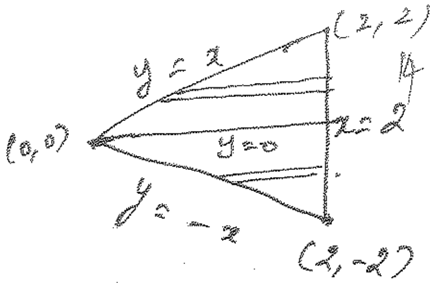
$$= \frac{1}{8}\int_{-x}^{x} x(x-y)\, dy$$

$$= \frac{1}{8}\int_{-x}^{x} (x^2 - xy)\, dy$$

$$= \frac{1}{8}\left(x^2 y - \frac{xy^2}{2}\right)_{-x}^{x} = \frac{1}{8}\left(\left(x^3 - \frac{x^3}{2}\right) - \left(-x^3 - \frac{x^3}{2}\right)\right)$$

$$= \frac{1}{8}(2x^3) = x^3/4$$

$$\therefore \boxed{f(x) = \frac{x^3}{4}, \quad 0 < x < 2}$$



$$f(y) = \begin{cases} \int_{x}^{2} f(x,y)\,dx, & 0 < y < 2. \\ \int_{-x}^{2} f(x,y)\,dx, & -2 < y < 0. \end{cases}$$

Case (i) $(0 < y < 2)$

$$f(y) = \int_{x}^{2} \frac{1}{8} x(x-y)\,dy = \frac{x}{8}\left(xy - \frac{y^2}{2}\right)_x^2$$

$$= \frac{x}{8}\left[(2x - 2) - \left(x^2 - \frac{x^2}{2}\right)\right]$$

$$= \frac{x}{8}\left(2x - 2 - \frac{x^2}{2}\right)$$

Case (i)

$$f(y) = \int_{y}^{2} \frac{1}{8} x(x-y)\,dx = \frac{1}{8}\int_{y}^{2}(x^2 - xy)\,dx$$

$$= \frac{1}{8}\left(\frac{x^3}{3} - \frac{x^2 y}{2}\right)_y^2 = \frac{1}{8}\left(\left(\frac{8}{3} - 2y\right) - \left(\frac{y^3}{3} - \frac{y^3}{2}\right)\right)$$

$$= \frac{1}{8}\left(\frac{8}{3} - 2y + \frac{y^3}{6}\right)$$

Case (ii)

$$f(y) = \int_{-y}^{2} \frac{1}{8} x(x-y)\,dx = \frac{1}{8}\int_{-y}^{2}(x^2 - xy)\,dx$$

$$= \frac{1}{8}\left(\frac{x^3}{3} - \frac{x^2 y}{2}\right)_{-y}^{2} = \frac{1}{8}\left\{\left(\frac{8}{3} - 2y\right) - \left(\frac{-y^3}{3} - \frac{y^3}{2}\right)\right\}$$

$$= \frac{1}{8}\left(\frac{8}{3} - 2y + \frac{5y^3}{6}\right)$$

$$\therefore\ f(y) = \begin{cases} \frac{1}{8}\left(\frac{8}{3} - 2y + \frac{y^3}{6}\right), & 0 < y < 2 \\ \frac{1}{8}\left(\frac{8}{3} - 2y + \frac{5y^3}{6}\right), & -2 < y < 0. \end{cases}$$

(c) $f(x/y)$ & $f(y/x)$

$$f(y/x) = \frac{f(x,y)}{f(x)} = \frac{\frac{1}{8}x(x-y)}{x^3/4} = \frac{x-y}{2x^2}$$

$$\therefore f(y/x) = \frac{x-y}{2x^2}, \quad -x < y < x$$

$$f(x/y) = \frac{f(x,y)}{f(y)} = \begin{cases} \dfrac{x(x-y)}{8/3 - 2y + \frac{y^3}{6}}, & y < x < 2 \\[2ex] \dfrac{x(x-y)}{8/3 - 2y + \frac{5y^3}{6}}, & -y < x < 2. \end{cases}$$

Sums for Practice

01) The two random variables X and Y have the joint density function $f(x,y) = 6(1-x-y)$, $x>0$, $y>0$, $x+y<1$. Find the marginal distribution of X and Y.

02) The joint density function of the random variables X and Y is given by $f(x,y) = \begin{cases} 8xy, & 0 \le x \le 1, 0 \le y \le \\ 0, & \text{otherwise}, \end{cases}$

Find the marginal and conditional distributions

03) Let X and Y have joint density function $f(x,y) = 2$, $0 < x < y < 1$. Find the marginal and conditional distributions

---

Problems based on Correlation and Regression

01) Two random variables X and Y have the following joint probability density function

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

Find the correlation co-efficient between X and Y.

Soln:

$$f(x) = \int_0^1 f(x, y)\, dy = \int_0^1 (2 - x - y)\, dy$$

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

$$\therefore \boxed{f(x) = \frac{3}{2} - x, \quad 0 \le x \le 1}$$

$$f(y) = \int_0^1 f(x, y)\, dx = \int_0^1 (2 - x - y)\, dx.$$

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

$$\therefore \boxed{f(y) = \frac{3}{2} - y, \quad 0 \le y \le 1}$$

$$E(X) = \int_0^1 x\, f(x)\, dx = \int_0^1 x\left(\frac{3}{2} - x\right) dx$$

$$= \left(\frac{3x}{2} - x^2\right)_0^1 = \left(\frac{3x^2}{4} - \frac{x^3}{3}\right)_0^1 = \frac{3}{4} - \frac{1}{3} = \frac{5}{12}$$

$$\therefore \boxed{E(X) = \frac{5}{12}}$$

$$E(Y) = \int_0^1 y f(y)\, dy = \int_0^1 y\left(\frac{3}{2} - y\right) dy = \int_0^1 \left(\frac{3y}{2} - y^2\right) dy$$

$$= \left(\frac{3y^2}{4} - \frac{y^3}{3}\right)_0^1 = \frac{3}{4} - \frac{1}{3} = \frac{5}{12}.$$

$$\therefore \boxed{E(Y) = \frac{5}{12}}$$

$$E(X^2) = \int_0^1 x^2 f(x)\, dx = \int_0^1 x^2\left(\frac{3}{2} - x\right) dx = \int_0^1 \left(\frac{3x^2}{2} - x^3\right) dx$$

$$= \left(\frac{3x^3}{6} - \frac{x^4}{4}\right)_0^1 = \frac{3}{6} - \frac{1}{4} = \frac{1}{4} \qquad \therefore \boxed{E(X^2) = \frac{1}{4}}$$

$$E(Y^2) = \int_0^1 y^2 f(y)\, dy = \int_0^1 y^2\left(\frac{3}{2} - y\right) dy = \int_0^1 \left(\frac{3y^2}{2} - y^3\right) dy$$

$$= \left(\frac{3y^3}{6} - \frac{y^4}{4}\right)_0^1 = \frac{3}{6} - \frac{1}{4} = \frac{1}{4} \qquad \therefore \boxed{E(Y^2) = \frac{1}{4}}$$

$$E(XY) = \int_0^1 \int_0^1 xy\, f(x,y)\, dx\, dy = \int_0^1 \int_0^1 xy(2 - x - y)\, dx\, dy$$

$$= \int_0^1 \int_0^1 (2xy - x^2y - xy^2)\, dx\, dy \quad \text{--- (1)}$$

$$\int_0^1 (2xy - x^2y - xy^2)\, dx = \left(\frac{2x^2y}{2} - \frac{x^3y}{3} - \frac{x^2y^2}{2}\right)_0^1$$

$$= y - \frac{y}{3} - \frac{y^2}{2} = \frac{2y}{3} - \frac{y^2}{2}.$$

$$\text{(1)} \Rightarrow E(xy) = \int_0^1 \left(\frac{2y}{3} - \frac{y^2}{2}\right) dy = \left(\frac{2y^2}{6} - \frac{y^3}{6}\right)_0^1$$

$$= \frac{2}{6} - \frac{1}{6} = \frac{1}{6}. \quad \therefore \boxed{E(xy) = \frac{1}{6}}$$

Now, $\operatorname{cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$

$$= \frac{1}{6} - \left(\frac{5}{12}\right)\left(\frac{5}{12}\right) = \frac{-1}{144}$$

$$\therefore \boxed{\operatorname{Cov}(X, Y) = \frac{-1}{144}}$$

$$\operatorname{Var} X = E(X^2) - [E(X)]^2 = \frac{1}{4} - \frac{25}{144} = \frac{11}{144}. \quad \sigma_X = \frac{\sqrt{11}}{12}$$

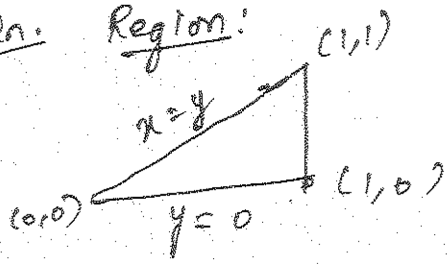
$$\operatorname{Var} Y = E(Y^2) - [E(Y)]^2 = \frac{1}{4} - \frac{25}{144} = \frac{11}{144}. \quad \sigma_Y = \frac{\sqrt{11}}{12}$$

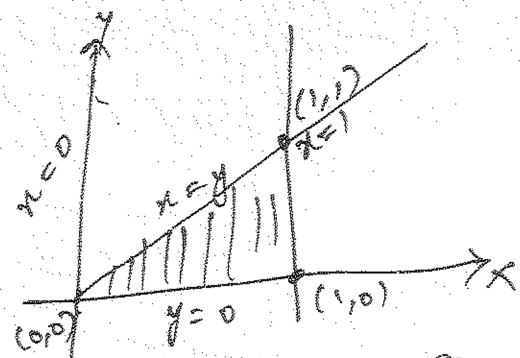
$$\operatorname{Corr}(X, Y) = \rho = \frac{\operatorname{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} = \frac{\frac{-1}{144}}{\frac{\sqrt{11}}{\cancel{12}} \cdot \frac{\sqrt{11}}{\cancel{12}}} = \frac{-1}{11}$$

$$\therefore \boxed{\operatorname{corr}(X, Y) = \frac{-1}{11}}$$

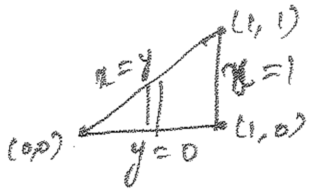
Q2) The joint density function of the Random variables X & Y is gn. by $f(x, y) = 8xy$, $0 \le x \le 1$, $0 \le y \le x$. Find the correlation & co-efficient & regression lines.

Soln. Region:



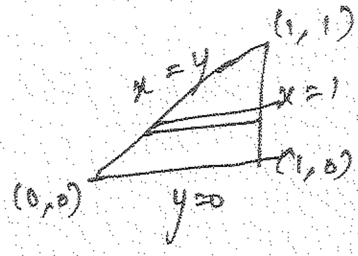


$$f(x) = \int_0^x f(x,y)\,dy = \int_0^x 8xy\,dy$$



$$= \overset{4}{\cancel{8}}x \left(\frac{y^2}{\cancel{2}}\right)_0^x = 4x(x^2) = 4x^3$$

$$\therefore \boxed{f(x) = 4x^3, \quad 0 < x < 1}$$



$$f(y) = \int_y^1 f(x,y)\,dx = \int_y^1 8xy\,dx$$

$$= {}^{4}\cancel{8}y \left(\frac{x^2}{\cancel{2}}\right)_y^1 = 4y(1-y^2)$$

$$\therefore \boxed{f(y) = 4y(1-y^2), \quad 0 < y < 1}$$

$$E(X) = \int_0^1 x\,f(x)\,dx = \int_0^1 x \cdot 4x^3\,dx = \int_0^1 4x^4\,dx$$

$$= \left(\frac{4x^5}{5}\right)_0^1 = \frac{4}{5} \qquad \therefore \boxed{E(X) = \frac{4}{5}}$$

$$E(X^2) = \int_0^1 x^2 f(x)\,dx = \int_0^1 x^2 \cdot 4x^3\,dx = \int_0^1 4x^5\,dx$$

$$= \left(\frac{4x^6}{6}\right)_0^1 = \frac{4}{6} \qquad \therefore \boxed{E(X^2) = 2/3}$$

$$E(Y) = \int_0^1 y \cdot f(y)\,dy = \int_0^1 y \cdot 4y(1-y^2)\,dy$$

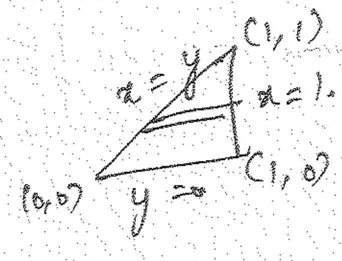
$$= \int_0^1 (4y^2 - 4y^4)\,dy = \left(\frac{4y^3}{3} - \frac{4y^5}{5}\right)_0^1 = 4/3 - \frac{4}{5} = 8/15$$

$$\therefore \boxed{E(Y) = \frac{8}{15}}$$

$$E(Y^2) = \int_0^1 y^2 f(y)\,dy = \int_0^1 y^2 \cdot 4y(1-y^2)\,dy$$

$$= \int_0^1 (4y^3 - 4y^5)\,dy = \left(y^4 - \frac{4y^6}{6}\right)_0^1 = \left(y^4 - \frac{2y^6}{3}\right)_0^1$$

$$= 1 - \frac{2}{3} = \frac{1}{3} \qquad \therefore \boxed{E(Y^2) = 1/3}$$



$$E(XY) = \int_0^1 \int_y^1 xy\, f(x,y)\,dx\,dy$$

$$= \int_0^1 \int_y^1 \cancel{xy}\; 8xy \cdot dx\,dy \qquad \text{——— (1)}$$

$$= \int_0^1 \int_y^1 8x^2y^2\,dx\,dy$$

$$\int_y^1 8x^2y^2\,dx = 8y^2\left(x^3/3\right)_y^1 = \frac{8y^2}{3}(1-y^3)$$

$$= \frac{8y^2}{3} - \frac{8y^5}{3}.$$

$$\therefore E(XY) = \int_0^1 \left(\frac{8y^2}{3} - \frac{8y^5}{3}\right) dy = \left(\frac{8y^3}{9} - \frac{8y^6}{18}\right)_0^1$$

$$= \left(\frac{8}{9} - \frac{8}{18}\right) - (0) = 4/9. \qquad \therefore \boxed{E(XY) = 4/9}$$

$$\therefore \operatorname{Cov}(X,Y) = E(XY) - E(X).E(Y)$$

$$= \frac{4}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right) = \frac{4}{225}.$$

$$\operatorname{Var}(X) = E(X^2) - [E(X)]^2 = \frac{2}{3} - \frac{16}{25} = \frac{2}{75}$$

$$\operatorname{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{1}{3} - \frac{64}{225} = \frac{11}{225}$$

$$\mathrm{Corr}(x,y) = \rho = \frac{\mathrm{cov}(x,y)}{\sqrt{\mathrm{var}\,x}\,\sqrt{\mathrm{var}\,y}} = \frac{4/225}{\sqrt{\frac{2}{75}}\,\sqrt{\frac{11}{225}}} = \frac{4}{\cancel{225}_{\cancel{15}\,3}} \cdot \frac{\cancel{5}\sqrt{3}}{\sqrt{2}} \cdot \frac{\cancel{15}}{\sqrt{11}}$$

$$= \frac{4\sqrt{3}}{3\sqrt{22}} = \frac{4}{\sqrt{66}}.$$

$$\therefore \boxed{\rho = 4/\sqrt{66}}$$

Regression line of x on y

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

$$\Rightarrow x - \frac{4}{5} = \frac{4}{\sqrt{66}} \cdot \frac{\sqrt{2}}{\sqrt{75}} \cdot \frac{\sqrt{225}}{\sqrt{11}}\left(y - \frac{8}{15}\right)$$

$$\Rightarrow \boxed{x - \frac{4}{5} = \frac{4}{11}\left(y - \frac{8}{15}\right)}$$

$$\frac{4}{\sqrt{66}} \cdot \frac{\sqrt{2}}{\cancel{5}\sqrt{3}} \cdot \frac{\cancel{15}^{\,3}}{\sqrt{11}}$$

$$= \frac{4 \cdot \cancel{3}^{\sqrt{3}} \cdot \sqrt{2}}{\sqrt{11}\,\sqrt{6}\,\sqrt{3}\,\sqrt{11}}$$

$$= \frac{4\sqrt{6}}{11\sqrt{6}} = \frac{4}{11}$$

Regression line of y on x

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

$$\Rightarrow y - \frac{8}{15} = \frac{4}{\sqrt{66}} \cdot \frac{\sqrt{11}}{\sqrt{225}} \cdot \frac{\sqrt{75}}{\sqrt{2}}\left(x - 4/5\right)$$

$$\Rightarrow \boxed{y - \frac{8}{15} = \frac{2}{3}\left(x - \frac{4}{5}\right)}$$

$$\frac{4}{\cancel{\sqrt{66}}_{\sqrt{6}}} \cdot \frac{\cancel{\sqrt{11}}}{\cancel{15}_{3}} \cdot \frac{\cancel{5}\sqrt{3}}{\sqrt{2}}$$

$$= \frac{4\sqrt{3}}{\cancel{\sqrt{6}}_{\sqrt{2}} \cdot 3\sqrt{2}}.$$

$$= \frac{\cancel{4}^{\,2}}{3 \cdot \cancel{2}} = 2/3$$

——— x ——— ——— x ———

03) Find the co-efficient of correlation from the following data.

X: 25 28 35 32 31 36 29 38 34 32

Y: 43 46 49 41 36 32 31 30 33 39.

Also find the regression lines. Find the value of Y when X = 30.

Soln:

x	y	x^2	y^2	xy
25	43	625	1849	1075
28	46	784	2116	1288
35	49	1225	2401	1715
32	41	1024	1681	1312
31	36	961	1296	1116
36	32	1296	1024	1152
29	31	841	961	899
38	30	1444	900	1140
34	33	1156	1089	1122
32	39	1024	1521	1248
$\Sigma x = 320$	$\Sigma y = 380$	10300 (Σx^2)	14838 (Σy^2)	12067 (Σxy)

$E(x) = \frac{1}{n}\Sigma x = \frac{1}{10}(320) = 32$

$E(y) = \frac{1}{n}\Sigma y = \frac{1}{10}(380) = 38$

$E(x^2) = \frac{1}{n}\Sigma x^2 = \frac{1}{10}(10380) = 1038$

$E(y^2) = \frac{1}{n}\Sigma y^2) = \frac{1}{10}(14838) = 1483.8$

$E(xy) = \frac{1}{n}\Sigma xy = \frac{1}{10}((12067) = 1206.7.$

$$\operatorname{cov}(x,y) = E(xy) - E(x)\cdot E(y)$$
$$= (1206.7) - (32)(38)$$
$$= -9.3.$$

$$\operatorname{Var}(x) = E(x^2) - [E(x)]^2 = 1038 - 1024 = 14$$

$$\operatorname{Var}(y) = E(y^2) - [E(y)]^2 = 1483.8 - 1444 = 39.8.$$

$$\operatorname{Corr}(x,y) = \frac{\operatorname{cov}(x,y)}{\sigma_x \cdot \sigma_y} = \frac{-9.3}{\sqrt{14}\sqrt{39.8}} = \frac{-9.3}{(3.74)(6.31)} = -0.39$$

$$\therefore \boxed{r = -0.39}$$

<u>Regression lines:</u>

<u>X on y:</u>

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

$$\Rightarrow x - 32 = (-0.39)\frac{(3.74)}{(6.31)}(y - 38).$$

$$\Rightarrow x - 32 = -0.23(y - 38).$$

$$\Rightarrow x - 32 = -0.23y + 8.74.$$

$$\Rightarrow \boxed{x + 0.23y = 40.74}$$

<u>Y on X</u>

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

$$\Rightarrow y - 38 = (-0.39)\frac{(6.31)}{3.74}(x - 32)$$

$$\Rightarrow y - 38 = -0.65(x - 32)$$

$\Rightarrow \quad y - 38 = -0.65x + 20.8$

$\Rightarrow \quad \boxed{0.65x + y = 58.8}$

To find the value of y when $X = 30$

To find y, we have to use the regression line of y on x.

$0.65x + y = 58.8$

$\Rightarrow \quad y = 58.8 - (0.65)(30)$

$= 39.3$

$\therefore \boxed{y = 39.3}$

Q4) The Joint probability mass function of X and Y is

X \ Y	-1	1
0	1/8	3/8
1	2/8	2/8

Find the correlation co-efft. of X and Y. Also find the regression lines.

Soln:

Marginal distribution of X

X	0	1
P(X = x)	1/2	1/2

Marginal distribution of Y

Y:	-1	1
P(Y=y)	3/8	5/8

$$E(x) = \sum x\, P(x=x) = 0\left(\tfrac{1}{2}\right) + (1)\left(\tfrac{1}{2}\right) = \tfrac{1}{2} \quad \therefore \boxed{E(x) = \tfrac{1}{2}}$$

$$E(x^2) = \sum x^2 P(x=x) = (0^2)\left(\tfrac{1}{2}\right) + (1)^2\left(\tfrac{1}{2}\right) = \tfrac{1}{2} \quad \therefore \boxed{E(x^2) = \tfrac{1}{2}}$$

$$E(y) = \sum y\, P(y=y) = (-1)\left(\tfrac{3}{8}\right) + (1)\left(\tfrac{5}{8}\right) = \tfrac{1}{4} \quad \boxed{E(y) = \tfrac{1}{4}}$$

$$E(y^2) = \sum y^2 P(y=y) = (-1)^2\left(\tfrac{3}{8}\right) + (1)^2\left(\tfrac{5}{8}\right) = 1 \quad \boxed{E(y^2) = 1}$$

$$E(xy) = \sum\sum xy\, P(x=x, y=y)$$

$$= (0)(-1)\left(\tfrac{1}{8}\right) + (0)(1)\left(\tfrac{3}{8}\right) + (1)(-1)\left(\tfrac{2}{8}\right) + (1)(1)\left(\tfrac{2}{8}\right)$$

$$= 0 + 0 - \tfrac{2}{8} + \tfrac{2}{8} = 0 \quad \therefore \boxed{E(xy) = 0}$$

$$\mathrm{Cov}(xy) = E(xy) - E(x)\cdot E(y) = 0 - \left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) = -\tfrac{1}{8}$$

$$\therefore \boxed{\mathrm{Cov}(xy) = -\tfrac{1}{8}}$$

$$\mathrm{Var}(x) = E(x^2) - [E(x)]^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \quad \boxed{\mathrm{var}(x) = \tfrac{1}{4}}$$

$$\mathrm{Var}(y) = E(y^2) - [E(y)]^2 = 1 - \frac{1}{16} = \frac{15}{16} \quad \boxed{\mathrm{var}(y) = \tfrac{15}{16}}$$

$$\mathrm{Corr}(x, y) = \frac{\mathrm{Cov}(xy)}{\sqrt{\mathrm{var}\,x}\,\sqrt{\mathrm{var}\,y}} = \frac{-\tfrac{1}{8}}{\tfrac{1}{2}\cdot\tfrac{\sqrt{15}}{4}} = -\frac{1}{\sqrt{15}}.$$

$$\therefore \boxed{\rho = -\tfrac{1}{\sqrt{15}}}$$

Regression lines

X on Y

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

$$\Rightarrow x - \frac{1}{2} = \left(\frac{-1}{\sqrt{15}}\right) \cdot \frac{1}{2} \cdot \frac{4}{\sqrt{15}} \cdot \left(y - \frac{1}{4}\right)$$

$$\Rightarrow \boxed{x - \tfrac{1}{2} = -\frac{2}{15}\left(y - \frac{1}{14}\right)}$$

Y on X

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

$$\Rightarrow y - \frac{1}{4} = \frac{-1}{\cancel{\sqrt{15}}} \cdot \frac{\cancel{\sqrt{15}}}{4} \cdot \frac{2}{1} \left(x - \tfrac{1}{2}\right)$$

$$\Rightarrow \boxed{y - \tfrac{1}{4} = -2\left(x - \tfrac{1}{2}\right)}$$

Sums for Practice

01) Calculate the correlation co-efficient between X & Y, regression lines for the following data.

X:	65	66	67	67	68	69	70	72
Y:	67	68	65	68	72	72	69	71

02) The joint density function of two random variables X & Y is $f(x,y) = \frac{8xy}{9}$, $1 < x < y < 2$.

Find the correlation co-efficient.

03) The joint density function of two random variables X & Y is $f(x,y) = \frac{1}{8}(6-x-y)$, $0 \le x \le 2$, $2 \le y \le 4$.

Find the correlation co-efficient.

04) The following table gives the joint probability distribution of two random variables X and Y. Find E(X), E(Y) and E(XY). Verify whether X and Y are correlated. [Hint: Show $Cov(X,Y) \neq 0$]

Y \ X	0	1	2	3
2	1/8	1/8	1/8	1/8
3	1/16	1/8	0	1/16
4	1/16	0	1/8	1/16

x x

More problems based on Correlation

01) Let X, Y and Z be uncorrelated random variables with zero means and standard deviations 5, 12 and 9 respectively. If $U = X+Y$ and $V = Y+Z$, find the correlation co-efficient between U and V.

Soln:

Given X, Y and Z are uncorrelated random variables.

(ie.) $Corr(X, Y) = 0 \Rightarrow Cov(X, Y) = 0$

$\Rightarrow E(XY) - E(X)E(Y) = 0$

$\Rightarrow \boxed{E(XY) = E(X)E(Y)}$ —— ①

|||ly $Corr(Y, Z) = 0 \Rightarrow \boxed{E(YZ) = E(Y)E(Z)}$ —— ②

$Corr(X, Z) = 0 \Rightarrow \boxed{E(XZ) = E(X)E(Z)}$ —— ③

Also, Given that $E(X) = 0,\ E(Y) = 0,\ E(Z) = 0$

$\sigma_X = 0,\ \sigma_Y = 0,\ \sigma_Z = 0$

①, ②, ③ $\Rightarrow \boxed{E(XY) = 0,\ E(YZ) = 0,\ E(XZ) = 0}$ —— ④

$\sigma_X = 5 \Rightarrow \sqrt{Var X} = 5 \Rightarrow Var X = 25$

$\Rightarrow E(X^2) - [E(X)]^2 = 25$

$\Rightarrow \boxed{E(X^2) = 25} \quad \because \boxed{E(X) = 0}$

|||ly $\sigma_Y = 12 \Rightarrow E(Y^2) = 144 \quad (Var Y = 144)$

$\sigma_Z = 9 \Rightarrow E(Z^2) = 81 \quad (Var Z = 81)$

$\therefore \boxed{E(X^2) = 25,\ E(Y^2) = 144,\ E(Z^2) = 81}$ —— ⑤

To find: $Corr(U, V)$ where $U = X + Y$, $V = Y + Z$

$$Corr(U, V) = \frac{Cov(UV)}{\sqrt{Var U}\sqrt{Var V}} \quad \text{—— ⑥}$$

$$\text{Cov}(U,V) = E(UV) - E(U)E(V)$$

$$= E\{(X+Y)(Y+Z)\} - E(X+Y)\,E(Y+Z)$$

$$= E\{XY + Y^2 + XZ + YZ\} - \{E(X) + E(Y)\}\{E(Y) + E(Z)\}$$

$$= E(XY) + E(Y^2) + E(XZ) + E(YZ) - 0 \quad (\text{from } 4)$$

$$= 144$$

$$\therefore \boxed{\text{Cov}(U,V) = 144}$$

$$\sqrt{\text{Var}\,U} = \sqrt{\text{Var}(X+Y)} = \sqrt{\text{Var}\,X + \text{Var}\,Y} = \sqrt{25+144} = 1$$

$$\sqrt{\text{Var}\,V} = \sqrt{\text{Var}(Y+Z)} = \sqrt{\text{Var}\,Y + \text{Var}\,Z} = \sqrt{144+81} = 15$$

$$\therefore (6) \Rightarrow \boxed{\text{corr}(U,V) = \frac{144}{(13)(15)} = \frac{48}{65}}$$

Q2) Let Z be a random variable with pdf $f(z) = \frac{1}{2}$, $-1 \leq z \leq 1$. Let the random variable $X = Z$ and $Y = Z^2$. Obviously X and Y are not independent since $X^2 = Y$. Show that X and Y are uncorrelated.

Soln: To prove X and Y are uncorrelated, we have to show $\text{Corr}(X,Y) = 0$

(ie) $\text{Cov}(X,Y) = 0$

Now $\quad cov(x, y) = E(xy) - E(x) \cdot E(y)$ —— (1) 22

$$E(xy) = E(z \cdot z^2) = E(z^3) = \int_{-1}^{1} z^3 f(z)\, dz$$

$$= \int_{-1}^{1} z^3 \left(\tfrac{1}{2}\right) dz = \frac{1}{2}\left(\frac{z^4}{4}\right)_{-1}^{1} = \frac{1}{2}\left(\tfrac{1}{4} - \tfrac{1}{4}\right) = 0$$

$\therefore \boxed{E(xy) = 0}$

$$E(x) = E(z) = \int_{-1}^{1} z \cdot f(z)\, dz = \int_{-1}^{1} z \left(\tfrac{1}{2}\right) dz$$

$$= \tfrac{1}{2}\left(\frac{z^2}{2}\right)_{-1}^{1} = \tfrac{1}{2}\left(\tfrac{1}{2} - \tfrac{1}{2}\right) = 0$$

$\therefore \boxed{E(x) = 0}$

$$E(y) = E(z^2) = \int_{-1}^{1} z^2 \cdot f(z)\, dz = \int_{-1}^{1} z^2 \left(\tfrac{1}{2}\right) dz$$

$$= \tfrac{1}{2}\left(\frac{z^3}{3}\right)_{-1}^{1} = \tfrac{1}{2}\left(\tfrac{1}{3} + \tfrac{1}{3}\right) = \tfrac{1}{2}\left(\tfrac{2}{3}\right) = \tfrac{1}{3}$$

$\therefore \boxed{E(y) = \tfrac{1}{3}}$

(1) $\Rightarrow \boxed{Cov(x, y) = 0}$

$\therefore$ x and y are uncorrelated

---

(03) If the random variables x and y have the variances 36 and 16 respectively, find the correlation co-efft. between $x+y$ and $x-y$.

Soln:

$$Corr(X+Y, X-Y) = \frac{Cov(X+Y, X-Y)}{\sqrt{Var(X+Y)}\ \sqrt{Var(X-Y)}} \quad \text{--- (1)}$$

Given : $\boxed{Var(X) = 36 \ \ \& \ \ var(Y) = 16}$ --- (2)

$$Cov(X+Y, X-Y) = E\{(X+Y)(X-Y)\} - E(X+Y).E(X-Y)$$

$$= E\{X^2 - Y^2\} - \{E(X) + E(Y)\}\{E(X) - E(Y)\}$$

$$= E(X^2) - E(Y^2) - \{[E(X)]^2 - [E(Y)]^2\}$$

$$= E(X^2) - E(Y^2) - [E(X)]^2 + [E(Y)]^2$$

$$= \{E(X^2) - [E(X)]^2\} - \{E(Y^2) - [E(Y)]^2\}$$

$$= var\,X - var\,Y = 36 - 16 = 20$$

$\therefore \boxed{Cov(X+Y, X-Y) = 20}$

$$\sqrt{var(X+Y)} = \sqrt{var\,X + var\,Y} = \sqrt{36+16} = \sqrt{52}$$

$$\sqrt{var(X-Y)} = \sqrt{var\,X + var\,Y} = \sqrt{36+16} = \sqrt{52}$$

$$(1) \Rightarrow Corr(X+Y, X-Y) = \frac{20}{\sqrt{52}\,\sqrt{52}} = \frac{20}{52} = 5/13$$

$\therefore \boxed{Corr(X+Y, X-Y) = \frac{5}{13}}$

---x---x---

05) Two variables X and Y have the same variance σ^2 and zero correlation. If $U = X + kY$ and $V = X + Y$, find the value of 'k' so that U and V are uncorrelated.

Soln:

Given: $\boxed{Var(X) = \sigma^2, \quad Var(Y) = \sigma^2}$

$Corr(X, Y) = 0 \Rightarrow Cov(X, Y) = 0$

$\Rightarrow E(XY) - E(X)E(Y) = 0$

$\Rightarrow \boxed{E(XY) = E(X).E(Y)}$

To find 'k' when $Corr(U, V) = 0$

$Corr(U, V) = 0 \Rightarrow Cov(UV) = 0$

$\Rightarrow E(UV) - E(U)E(V) = 0$

$\Rightarrow E(UV) = E(U)E(V)$

$\Rightarrow E\{(X+kY)(X+Y)\} = E(X+kY)\,E(X+Y)$

$\Rightarrow E\{X^2 + kXY + XY + kY^2\} = \{E(X) + kE(Y)\}\{E(X) + E(Y)\}$

$\Rightarrow E(X^2) + \cancel{kE(XY)} + \cancel{E(XY)} + kE(Y^2)$

$= [E(X)]^2 + \cancel{k\,E(X)E(Y)} + \cancel{E(X)E(Y)} + k[E(Y)]^2$

$(\because E(XY) = E(X)E(Y))$

$\Rightarrow E(X^2) - [E(X)]^2 + k\{E(Y^2) - [E(Y)]^2\} = 0$

$\Rightarrow Var\,X + k\,Var\,Y = 0$

$\Rightarrow \quad \sigma^2 + k\sigma^2 = 0$

$\Rightarrow \quad k\sigma^2 = -\sigma^2$

$\Rightarrow \quad \boxed{k = -1}$

<u>Sums for Practice</u>

01) Two variables X and Y have the same variance σ^2 and zero correlation. Show that $U = X\cos\alpha + Y\sin\alpha$, $V = X\sin\alpha - Y\cos\alpha$ have the same variance σ^2 and zero correlation.

02) Find an expression for the correlation co-effts of U and V if $U = X + Y$ and $V = X - Y$.

03) The following data were available $\bar{x} = 970$, $\bar{y} = 18$, $\sigma_x = 38$, $\sigma_y = 2$, Correlation co-efficient $r = 0.6$. Find the line of regression and obtain the value of x given $y = 20$.

04). Two random variables X and Y are defined as $Y = 4X + 9$. Find the correlation co-effts between X and Y.

Problems based on regression lines

Q1) The two lines of regression are $8x - 10y + 66 = 0$; $40x - 18y - 214 = 0$. The variance of x is 9. Find (i) the mean values of x and y (ii) correlation co-effts between x and y and (iii) standard deviation of y.

Soln:

(i) Mean values of x and y

∵ the regression lines pass through $(\bar{x}, \bar{y})$, we have

$$8\bar{x} - 10\bar{y} = -66 \quad \text{—} \; (1)$$

$$40\bar{x} - 18\bar{y} = +214 \quad \text{—} \; (2)$$

$$(1) \times 5 \Rightarrow \quad 40\bar{x} - 50\bar{y} = -330$$

$$(2) \Rightarrow \quad \overset{(-)}{40\bar{x}} \overset{(+)}{- 18\bar{y}} = \overset{(-)}{+214}$$

$$-32\bar{y} = -544 \Rightarrow \boxed{\bar{y} = 17}$$

$$(1) \rightarrow \quad 8\bar{x} - 170 = -66 \Rightarrow 8\bar{x} = 104 \Rightarrow \boxed{\bar{x} = 13}$$

$$\therefore \boxed{\bar{x} = 13, \; \bar{y} = 17}$$

(ii) Correlation Co-efficient between x and y

Let us assume x on y be $8x - 10y + 66 = 0$ — (3)

y on x be $40x - 18y - 214 = 0$ — (4)

③ $\Rightarrow$ $8x = 10y - 66$

$\Rightarrow x = \frac{10}{8}y - \frac{66}{8}$

$\therefore \boxed{r\frac{\sigma_x}{\sigma_y} = \frac{10}{8}}$

④ $\Rightarrow$ $18y = 40x - 214 \Rightarrow y = \frac{40}{18}x - \frac{214}{18}$

$\therefore \boxed{r\frac{\sigma_y}{\sigma_x} = \frac{40}{18}}$

Now $r\frac{\sigma_x}{\sigma_y} \times r\frac{\sigma_y}{\sigma_x} = \frac{\cancel{10}^{5}}{\cancel{8}_{4}} \times \frac{\cancel{40}^{\cancel{10}^{5}}}{\cancel{18}_{9}} \Rightarrow r^2 = 25/9$

$\therefore r = \pm 5/3 \Rightarrow r = \pm 1.6$ which is impossible.

$\therefore$ Our assumption is wrong.

$\therefore$ x on y : $40x - 18y - 214 = 0 \Rightarrow \boxed{r\frac{\sigma_x}{\sigma_y} = \frac{18}{40}}$

y on x : $8x - 10y + 66 = 0 \rightarrow \boxed{r\frac{\sigma_y}{\sigma_x} = \frac{8}{10}}$

$\therefore r\frac{\sigma_x}{\sigma_y} \cdot r\frac{\sigma_y}{\sigma_x} = \frac{18}{40} \cdot \frac{8}{10} \Rightarrow r^2 = 9/25$

$\therefore r = \pm 3/5 \Rightarrow \boxed{r = \pm 0.6}$

(iii) Standard deviation of y

Given $\text{Var } x = 9$

$\Rightarrow$ S.D of $x = 3$ ie $\boxed{\sigma_x = 3}$

$\because \; r \dfrac{\sigma_x}{\sigma_y} = \dfrac{18}{40} \Rightarrow \dfrac{\cancel{3}}{\cancel{8}} \cdot \dfrac{\cancel{3}}{\sigma_y} = \dfrac{\cancel{18}^{2}}{\cancel{40}_{8}}$

$\Rightarrow \dfrac{1}{\sigma_y} = \dfrac{1}{4} \Rightarrow \boxed{\sigma_y = 4}$

Sums for Practice

01) The regression equation of x on y is $3y - 5x + 108 = 0$. If the mean value of y is 44 and the variance of x were $\frac{9}{16}$ th of the variance of y. Find the mean value of x and the correlation co-efficient.

02) Can $y = 5 + 2.8x$ and $x = 3 - 0.5y$ be the estimated regression equations of y on x and x on y respectively? Explain your answer.

03) A statistical investigation obtain the following regression equation in survey $x - y - 6 = 0$ and $0.64x - 4.08 = 0$
Here x = age of husband & y = age of wife.
Find (i) Mean of x and y
(ii) Correlation Co-eff between x and y
(iii) σ_y if $\sigma_x = 4$.

Transformation of Random Variables

Here we transform a two dimensional random variable (X, Y) into another two dimensional random variable (U, V). where U and V are defined by the eqns of the form $U = f_1(X, Y)$ & $V = f_2(X, Y)$.

Procedure:

1. Find the joint pdf of (X, Y) if it is not given
2. $U = f_1(X, Y)$ & $V = f_2(X, Y)$ will be given.

   From this we have to express the relation in the form $X = g_1(U, V)$, $Y = g_2(U, V)$.
3. Find $\dfrac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix}$
4. Find the joint pdf of (U, V) using the formula

   $f(u, v) = \left|\dfrac{\partial(x, y)}{\partial(u, v)}\right| f(x, y)$ (Here $f(x, y)$ should be written in terms of u & v using step 2).
5. Find the marginal densities $f(u)$ & $f(v)$ from $f(u, v)$.
6. Find the limits for $f(u, v)$, $f(u)$ & $f(v)$ using the relation in step 2.

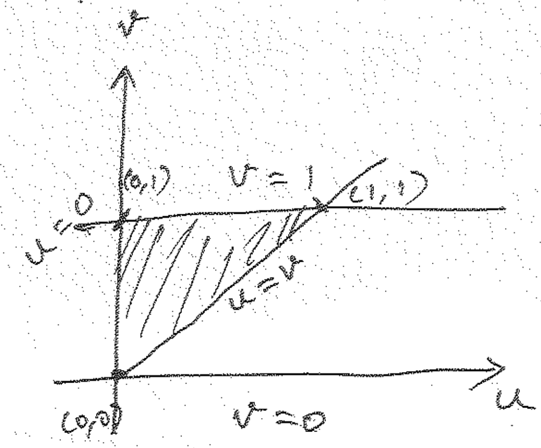
To find the limits for $U + V$

$\because \; 0 < x < 1 \;,\; 0 < y < 1$

$\Rightarrow \; 0 < \frac{U}{V} < 1 \;,\; 0 < V < 1$

$\Rightarrow \; 0 < u < v \;,\; 0 < v < 1$

$\therefore \; \boxed{f(u, v) = 1, \quad 0 < u < v, \quad 0 < v < 1}$

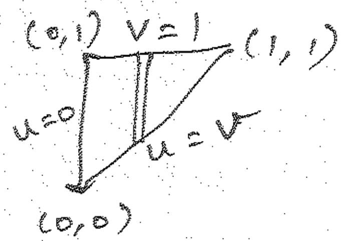


To find $f(u)$

$$f(u) = \int_u^1 f(u, v)\, dv$$

$$= \int_u^1 1 \,.\, dv$$

$$= (v)_u^1 = 1 - u$$



$\therefore \; \boxed{f(u) = 1 - u, \quad 0 < u < 1}$

Q2) The joint pdf of X and Y is given by $f(x, y) = e^{-(x+y)}$, $x > 0, y > 0$, find the pdf of $U = \frac{X+Y}{2}$

Soln: $f(x, y) = e^{-(x+y)}$, $x > 0, y > 0$

Here $U = \frac{X+Y}{2}$, $V = Y$

Q1) If X and Y are independent variates uniformly distributed in (0, 1), find the distribution of XY.

Soln:

Let $U = XY$ and $V = Y$.

Given X follows uniform distribution in $(0, 1)$

$\therefore f(x) = 1, \quad 0 \le x \le 1.$ $\left(f(x) = \frac{1}{b-a}, a < x < b \right)$

|||ly $f(y) = 1, \quad 0 \le y \le 1$

$\because$ X and Y are independent,

$f(x, y) = f(x) \cdot f(y)$

$= 1$

$\therefore \boxed{f(x, y) = 1, \quad 0 \le x \le 1, \quad 0 \le y \le 1}$

$\because U = XY$ & $V = Y$

Now $Y = V \Rightarrow X = \frac{U}{Y} = \frac{U}{V}$

$\therefore \boxed{Y = V}, \quad \boxed{X = \frac{U}{V}}$

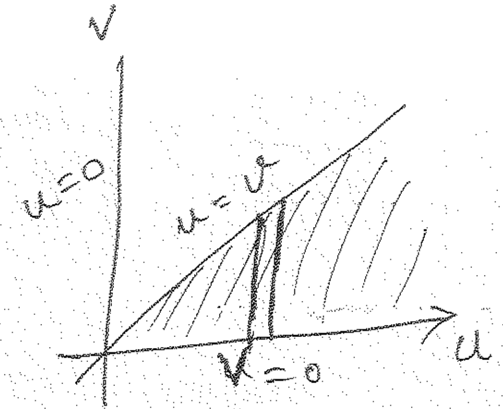
Now $\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} 1/v & -u/v^2 \\ 0 & 1 \end{vmatrix} = \frac{1}{v}.$

$\therefore f(u, v) = |J| \, f(x, y)$

$= \frac{1}{v}$

$$\therefore \boxed{f(u,v) = 6e^{-3u+v}, \quad u>0,\ v>0,\ u>v}$$

To find $f(u)$



$$f(u) = \int_0^u f(u,v)\,dv$$

$$= \int_0^u 6e^{-3u+v}\,dv$$

$$= \left(6e^{-3u+v}\right)_0^u = 6e^{-3u+u} - 6e^{-3u}$$

$$= 6e^{-2u} - 6e^{-3u}$$

$$\therefore \boxed{f(u) = 6e^{-2u} - 6e^{-3u}, \quad 0<u<\infty}$$

---

04) If the joint density of X_1 and X_2 is given by $f(x_1, x_2) = e^{-(x_1+x_2)}$, $x_1, x_2 > 0$, find the pdf of

$$y = \frac{x_1}{x_1+x_2}.$$

Soln: For convenience. $x_1 = x$, $x_2 = y$

$$\therefore f(x,y) = e^{-(x+y)}, \quad x>0,\ y>0$$

Given $U = \dfrac{X}{X+Y}$, Let $V = X+Y$

(we took $Y = U$)

Now $U = \dfrac{X}{V} \Rightarrow \boxed{X = uv}$

Also $V = x + y \Rightarrow V = UV + y$

$\Rightarrow \boxed{y = V - UV}$

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v - uv + uv$$
$$= v$$

$$\therefore f(u, v) = |J| \cdot f(x, y) = v \cdot e^{-(x+y)} = v e^{-(\cancel{uv} + v - \cancel{uv})}$$

$$\therefore \boxed{f(u, v) = v e^{-v}}$$

Limits for U & V

$x > 0, \ y > 0$

$\Rightarrow uv > 0 \Rightarrow \boxed{u > 0, \ v > 0}$

$y > 0 \Rightarrow v - uv > 0 \Rightarrow v > uv \Rightarrow 1 > u$

$\therefore \boxed{u < 1}$

$$\therefore \boxed{f(u, v) = v e^{-v}, \ 0 < u < 1, \ v > 0}$$

To find $f(u)$

$$f(u) = \int_0^\infty f(u, v)\, dv = \int_0^\infty v e^{-v}\, dv$$

$$= \left\{ v(-e^{-v}) - (1)(e^{-v}) \right\}_0^\infty = (0) - (-1) = 1$$

$$\therefore \boxed{f(u) = 1, \ 0 < u < 1}$$

05) Let (X, Y) be a two dimensional non-negative continuous random variable having the joint density $f(x,y) = 4xy e^{-(x^2+y^2)}$, $x>0, y>0$ find the density function of $U = \sqrt{X^2+Y^2}$.

Soln: Given $f(x,y) = 4xy e^{-(x^2+y^2)}$, $x>0, y>0$

$U = \sqrt{X^2+Y^2}$, $\boxed{Y = V}$

Now $X^2+Y^2 = U^2 \Rightarrow X^2 = U^2 - Y^2 = U^2 - V^2$

$\therefore \boxed{X = \sqrt{u^2-v^2}}$

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} \frac{1}{\cancel{2}\sqrt{u^2-v^2}}(\cancel{2}u) & \frac{1}{2\sqrt{u^2-v^2}}(-2v) \\ 0 & 1 \end{vmatrix}$$

$$= \frac{u}{\sqrt{u^2-v^2}}.$$

$$\therefore f(u,v) = |J|\, f(x,y) = \frac{u}{\sqrt{u^2-v^2}} 4xy e^{-(x^2+y^2)}$$

$$= \frac{u}{\cancel{\sqrt{u^2-v^2}}} 4 \cdot \cancel{\sqrt{u^2-v^2}}\,(v)\, e^{-u^2}$$

$$\boxed{f(u,v) = 4uv e^{-u^2}}$$

Limits for U & V

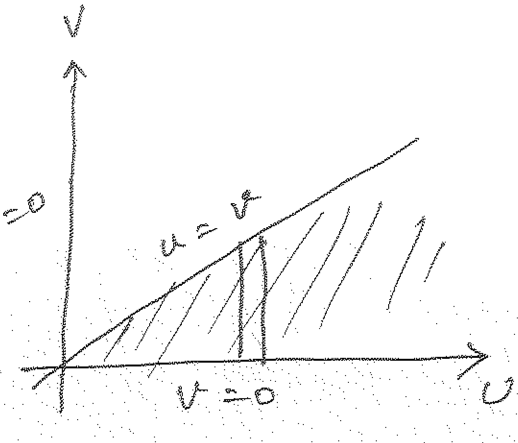
$x>0 \Rightarrow \sqrt{u^2-v^2} > 0 \Rightarrow u^2-v^2>0 \Rightarrow u^2>v^2$

$\Rightarrow \boxed{u>v}$

$y>0 \Rightarrow \boxed{v>0}$ Also $u>0$, $\because x>0, y>0, u=\sqrt{x^2+y^2}$

To find $f(u)$

$$f(u) = \int_0^u f(u,v)\,dv$$



$$= \int_0^u 4uv\,e^{-u^2}\,dv$$

$$= 4ue^{-u^2}\left(\frac{v^2}{2}\right)_0^u = 4ue^{-u^2}\left(\frac{u^2}{2}\right)$$

$$= 2u^3e^{-u^2}$$

$$\therefore \boxed{f(u) = 2u^3e^{-u^2}, \quad u > 0}$$

Sums for Practice

01) If X and Y are independent exponential random variables each with parameter 1, find the pdf of $U = X - Y$.

02) If X and Y are independent variables uniformly distributed in $(0,1)$, find the distribution of $\frac{X}{Y}$.

03) If the joint pdf of X and Y is $f(x,y) = x+y$, $0 \le x, y \le 1$. find the pdf of $U = XY$.

04) If X and Y are independent R.V.s with pdf's e^{-x}, $x \geq 0$ and e^{-y}, $y \geq 0$ respectively, P.T $U = \frac{X}{X+Y}$ and $V = X+Y$ are independent

---

Central Limit theorem:

Statement

If $X_1, X_2, \ldots, X_n$ is a sequence of independent identically distributed random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$, $i = 1, 2, 3, \ldots$ and

(i) if $S_n = X_1 + X_2 + \cdots + X_n$, then S_n follows normal distribution with mean $n\mu$ and variance $n\sigma^2$.

(ie) $S_n \sim N(n\mu, \sigma\sqrt{n})$.

(ii) if $\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$, then $\bar{X}$ follows normal distribution with mean μ and variance σ^2/n.

(ie) $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$.

Problems

01) If $X_1, X_2, \ldots, X_n$ are Poisson variables with parameter $\lambda = 2$, use the central limit theorem to evaluate $P(120 < S_n < 160)$ where $S_n = X_1 + X_2 + \cdots + X_n$ & $n = 75$.

Soln: $\because$ $X_1, X_2, \ldots, X_n$ follows Poisson Distribution,

$E(X_i) = \lambda$ & $Var(X_i) = \lambda$

$\therefore$ $E(X_i) = 2 \Rightarrow \boxed{\mu = 2}$

$Var(X_i) = 2 \Rightarrow \boxed{\sigma^2 = 2}$

Given, $\boxed{n = 75}$

$\therefore S_n \sim N(n\mu, \sigma\sqrt{n})$

$\Rightarrow S_n \sim N(150, \sqrt{150})$

<u>To find $P(120 < S_n < 160)$</u>

Here $z = \dfrac{S_n - 150}{\sqrt{150}}$

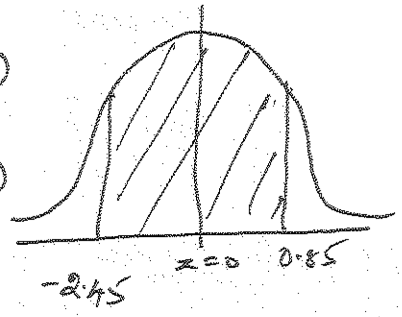
If $S_n = 120$, $z = \dfrac{120 - 150}{\sqrt{150}} = -2.45$

If $S_n = 160$, $z = \dfrac{160 - 150}{\sqrt{150}} = 0.85$

$\therefore P(120 \le S_n \le 160) = P(-2.45 \le z \le 0.85)$

$= P(-2.45 \le z \le 0) + P(0 \le z \le 0.85)$

$= P(0 \le z \le 2.45) + P(0 \le z \le 0.85)$



$= 0.4927 + 0.2939$

$= 0.7866$

$\therefore \boxed{P(120 < S_n < 160) = 0.7866}$

---

02) The resistors r_1, r_2, r_3 and r_4 are independent random variables and is uniform in the interval $(450, 550)$. Using the central limit theorem, find $P(1900 \le r_1 + r_2 + r_3 + r_4 \le 2100)$.

Soln:

$\because r_1, r_2, r_3, r_4$ follows uniform distribution

$$E(r_i) = \frac{a+b}{2} \quad \& \quad var(r_i) = \frac{(b-a)^2}{12}$$

(Here $a = 450$, $b = 550$)

$\therefore E(r_i) = 500$, $var(r_i) = 833.33$.

$\therefore \boxed{\mu = 500}$, $\boxed{\sigma^2 = 833.33}$ $\boxed{n = 4}$

$\therefore S_n \sim N(n\mu, \sigma\sqrt{n})$

$\Rightarrow S_n \sim N(2000, \sqrt{3333.2}) \Rightarrow S_n \sim N(2000, 57.73)$

To find $P(1900 \le S_n \le 2100)$

Here $z = \dfrac{S_n - 2000}{57.73}$

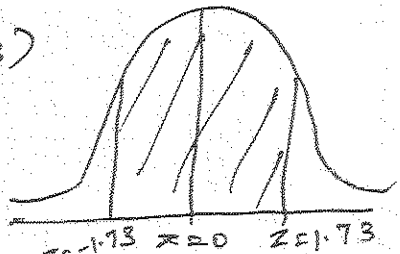
If $S_n = 1900$, $z = \dfrac{1900 - 2000}{57.73} = -1.73$

If $S_n = 2100$, $z = \dfrac{2100 - 2000}{57.73} = 1.73$

$P(1900 \le S_n \le 2100) = P(-1.73 \le z \le 1.73)$

$= P(-1.73 \le z \le 0) + P(0 \le z \le 1.73)$

$= P(0 \le z \le 1.73) + P(0 \le z \le 1.73)$



$= 0.4582 + 0.4582$

$= 0.9164$

$\therefore \boxed{P(1900 \le S_n \le 2100) = 0.9164}$

Q3) A coin is tossed 300 times. What is the probability that heads will appear more than 140 times and less than 150 times.

<u>Soln:</u>

p — prob. of getting head $= 1/2$

$\Rightarrow q = 1/2$ Also $n = 300$ ($n \rightarrow$ no. of trials)

Here the r.v follows Binomial Distribution

$E(X_i) = np = 150.$

$Var(X_i) = npq = 75.$

$\therefore \boxed{\mu = 150, \; \sigma^2 = 75, \; n = 1}$ (no. of random variables)

$\therefore S_n \sim N(n\mu, \sigma\sqrt{n})$

$\Rightarrow S_n \sim N(150, \sqrt{75})$

<u>To find $P(140 < S_n < 150)$</u>

Here $Z = \dfrac{S_n - 150}{\sqrt{75}}$

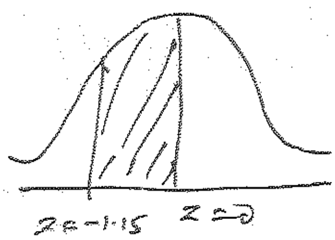
If $S_n = 140$, $Z = \dfrac{140 - 150}{\sqrt{75}} = -1.15$

If $S_n = 150$, $Z = \dfrac{150 - 150}{\sqrt{75}} = 0.$

$\therefore P(140 \leq S_n \leq 150) = P(-1.15 \leq Z \leq 0)$

$= P(0 \leq Z \leq 1.15)$

$= 0.3749.$



$\therefore \boxed{P(140 \leq S_n \leq 150) = 0.3749}$

04) The life time of a certain brand of a Tube light may be considered as a Random variable with mean 1200 hrs and standard deviation 250 hrs. Find the probability, using central limit theorem, that the average life time of 60 lights exceeds 1250 hrs.

<u>Soln</u>:

Given $E(X_i) = 1200$ & $Var(X_i) = (250)^2$.

$$\Rightarrow \boxed{\mu = 1200, \quad \sigma^2 = (250)^2, \quad n = 60}$$

$\bar{X} \rightarrow$ Mean life time of 60 lights.

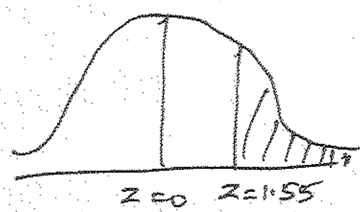
$\because \bar{X} \sim N(\mu, \sigma/\sqrt{n})$

$\Rightarrow \bar{X} \sim N\left(1200, \dfrac{250}{\sqrt{60}}\right)$

<u>To find $P(\bar{X} > 1250)$</u>

Here $z = \dfrac{\bar{X} - 1200}{250/\sqrt{60}}$

If $\bar{X} = 1250$, $z = \left(\dfrac{1250 - 1200}{250}\right)\sqrt{60} = 1.55$



$\therefore P(\bar{X} > 1250) = P(z > 1.55)$

$= P(0 < z < \infty) - P(0 < z < 1.55)$

$= 0.5 - 0.4394$

$= 0.0606.$

$$\therefore \boxed{P(\bar{X} > 1250) = 0.0606}$$

05) A random sample of size 100 is taken from a population whose mean is 60 and variance is 400. Using central limit theorem, with what probability can we assert that the mean of the sample will not differ from $\mu = 60$ by more than 4.

Soln:

Given $\mu = 60$, $\sigma^2 = 400$, $n = 100$.

$\therefore \bar{x} \sim N(\mu, \sigma^2/\sqrt{n})$

$\Rightarrow \bar{x} \sim N(60, 2)$

To find $P(\text{diff. between } \bar{x} \text{ \& } \mu = 60 < 4)$

(ie) $P(|\bar{x} - \mu| < 4)$

$= P(|\bar{x} - 60| < 4)$

$= P(-4 \le \bar{x} - 60 \le 4)$

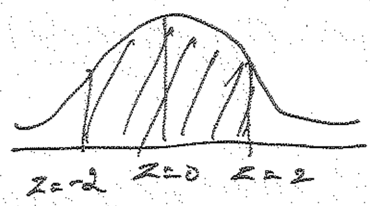
$= P(56 \le \bar{x} \le 64)$.

(ie) $P(56 \le \bar{x} \le 64)$

Here $z = \dfrac{\bar{x} - 60}{2}$.

If $\bar{x} = 56$, $z = \dfrac{56-60}{2} = -2$

If $\bar{x} = 64$, $z = \dfrac{64-60}{2} = 2$.



$\therefore P(56 \le \bar{x} \le 64) = P(-2 \le z \le 2)$

$= P(-2 \le z \le 0) + P(0 \le z \le 2)$

$= P(0 \le z \le 2) + P(0 \le z \le 2)$

$= 0.4773 + 0.4773 = 0.9546$

$\therefore \boxed{P(56 < \bar{x} < 64) = 0.9546}$

———— x ———— x ————

06) A distribution with unknown mean μ has variance equal to 1.5. Use Central limit theorem to determine how large a sample should be taken from the distribution in order that the probability will be atleast 0.95 that the sample mean will be within 0.5 of the population mean.

Soln:

Given: $\boxed{\sigma^2 = 1.5, \quad \mu = \mu. \quad n = ?}$

$\because \bar{x} \sim N(\mu, \sigma/\sqrt{n})$

$\Rightarrow \bar{x} \sim N(\mu, \frac{\sqrt{1.5}}{\sqrt{n}})$.

Given that $P(|\bar{x} - \mu| < 0.5) \geq 0.95$

$\Rightarrow P(-0.5 < \bar{x} - \mu < 0.5) \geq 0.95$

$\Rightarrow P(\mu - 0.5 < \bar{x} < \mu + 0.5) \geq 0.95$

$\because z = \frac{\bar{x} - \mu}{\sqrt{1.5}/\sqrt{n}}$.

If $\bar{x} = \mu - 0.5$, $z = \frac{\mu - 0.5 - \mu}{\sqrt{1.5}/\sqrt{n}} = -0.4082\sqrt{n}$.

If $\bar{x} = \mu + 0.5$, $z = \frac{\mu + 0.5 - \mu}{\sqrt{1.5}/\sqrt{n}} = 0.4082\sqrt{n}$.

$\therefore P(\mu - 0.5 < \bar{x} < \mu + 0.5) \geq 0.95$

$\Rightarrow P(-0.4082\sqrt{n} < z < 0.4082\sqrt{n}) \geq 0.95$

$$\Rightarrow P(-0.4082\sqrt{n} < z < 0) + P(0 < z < 0.4082\sqrt{n}) \geq 0.95$$

$$\Rightarrow P(0 < z < 0.4082\sqrt{n}) + P(0 < z < 0.4082\sqrt{n}) \geq 0.95$$

$$\Rightarrow 2P(0 < z < 0.4082\sqrt{n}) \geq 0.95$$

$$\Rightarrow P(0 < z < 0.4082\sqrt{n}) \geq 0.475 \quad \text{--- (1)}$$

In normal tables.

$$P(0 < z < 1.96) = 0.475 \quad \text{--- (2)}$$

From (1) & (2),

$$0.4082\sqrt{n} > 1.96$$

$$\Rightarrow \sqrt{n} > \frac{1.96}{0.4082}$$

$$\Rightarrow n > 24.$$

$\therefore$ The size of the sample must be atleast 24.

Sums for Practice

01) If X_i, $i = 1, 2, \ldots 50$ are independent R.V.s each having a poisson distribution with parameter $\lambda = 0.03$ and $S_n = X_1 + X_2 + \ldots + X_n$, evaluate $P(S_n \geq 3)$.

02) The guaranteed average life of a certain type of electric light bulb is 1000 hrs with a standard