Unit 3: Random variables

- Discrete Random variables
- Continous Random variables
- Discrete joint probability distribution
- Continuous joint probability distribution

Random Variable:

A random variable ("r.v.") associates a value (a number) to every possible outcome of a random experiment.

Mathematically:

It is a function from the sample space Ω to the real numbers, $\mathbb{R}$.

i.e $X : \Omega \longrightarrow \mathbb{R}$

Notation:

random variable X, Y, so on (upper case),

numerical value x, y, so on (corresponding lower case)

Ex: Toss a coin three times. Let X be the random variable which denote number of tails.

Sample space,

$\Omega = \{HHH, THH, HTH, HHT, TTH, THT, HTT, TTT\}$

and r.v.

$X = \{0, 1, 2, 3\}$

Here X takes discrete value, $\therefore$ it is called discrete r.v.

Ex: We define X as the height of randomly selected individuals.

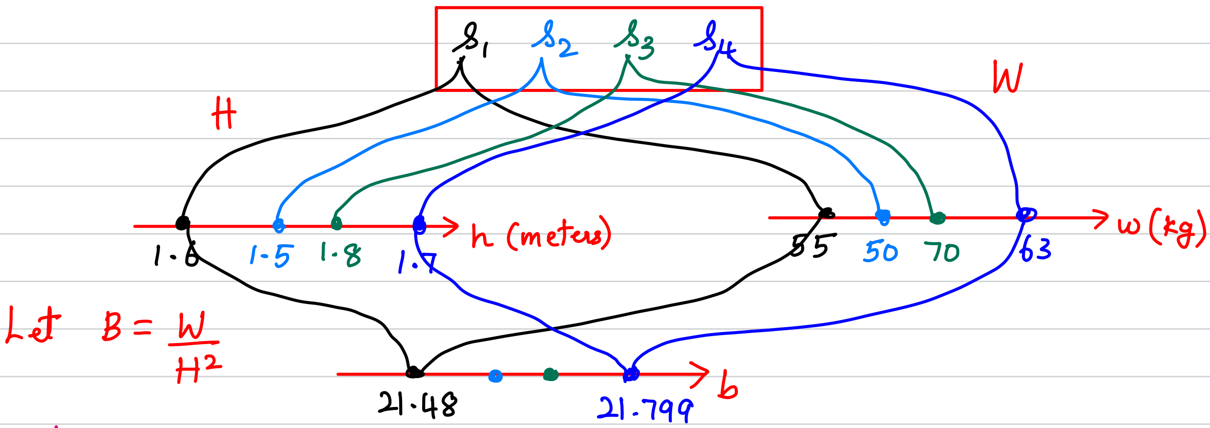
Height is measured continuously. $\therefore$ Here X takes continuous values in an interval. It is called continuous r.v

That is, random variables can take discrete or continuous values accordingly they are called discrete r.v. or continuous r.v.

- A r.v. that can take on at most a countable no. of possible values is said to be discrete.
- A r.v that takes uncountable possible values is called continuous.

Ex: Let sample space Ω be set of students. For instance, let $\Omega = \{s_1, s_2, s_3, s_4\}$

Let H be a r.v which define height of students and let W be a r.v which define weight of students



Note:

1) We can have several random variables defined on the same sample space.

2) A function of one or several random variables is also a random variable.

If X and Y are random variables, then $X+Y$ is also a r.v. It takes value $x+y$ when X takes x and Y takes y.

Discrete random Variable X

Defn: Probability mass function: [PMF]

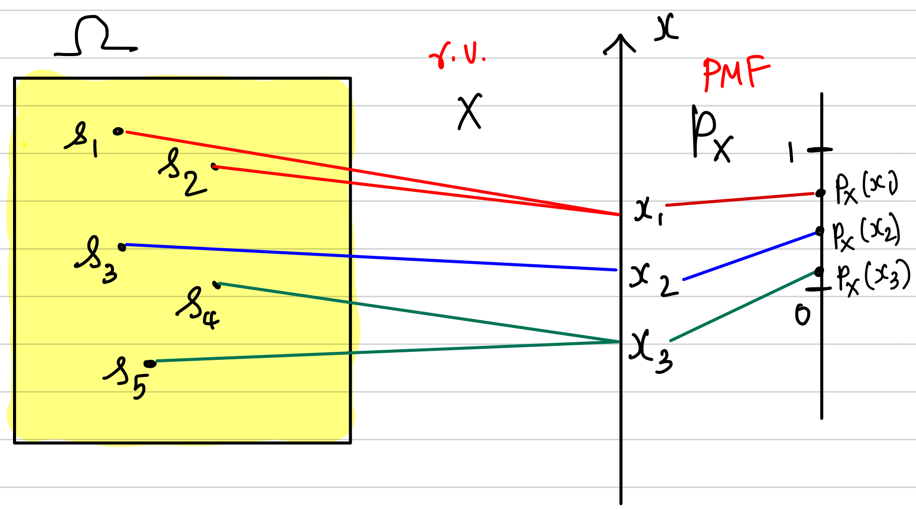
For a discrete r.v. X, we can define a function that assigns to every x in X a probability of its occurence; it is called the Probability mass function.

It is denoted by

$$p_X(x) = P(X=x) = P(\{\omega \in \Omega \text{ such that } X(\omega)=x\}).$$

- If Y is another r.v, then its PMF is denoted by $p_Y(y)$

For instance,



Here

$P[X=x_1) = P(s_1) + P(s_2)$

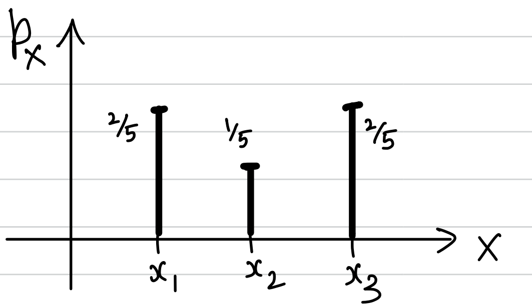
$P(X=x_2) = P(s_3)$

$P(X=x_3) = P(s_4) + P(s_5)$

If every element Ω are equally likely to occur, then

$P(X=x_1) = 2/5, \quad P(X=x_2) = 1/5, \quad P(X=x_3) = 2/5$

Probability bar chart



Properties:

$$p_X(x) \geq 0 \quad \text{and} \quad \sum_x p_X(x) = 1.$$

Note:

- If $f(x)$ is a PMF defined on r.v X, then
  $f(x) \geq 0$ and $\sum_x f(x) = 1$
- PMF is the "probability law" or "probability distribution" of X.

Cumulative distribution function (cdf)

If $p_X(x)$ is a probability distribution of r.v. X, then the cumulative distribution function or just distribution function, $F_X(x)$ of the random variable is

$$F_X(x) = P(X \leq x) = \sum_{\leq x} p_X(x)$$

$F_X(x)$ satisfies the following properties

1) $0 \leq F_X(x) \leq 1$
2) If $x_1 \leq x_2$ then $F_X(x_1) \leq F_X(x_2)$
3) If $X = \{x_1, x_2, x_3, \ldots, x_k\}$, then

$$P(X = x_i) = F_X(x_i) - F_X(x_{i-1}), \quad i = 2, 3, \ldots, k$$

$$P(X = x_1) = F_X(x_1)$$

Ex: PMF Calculation.

Two rolls of a six faced die.

$Y = 2^{nd}$ roll						
6	7	8	9	10	11	12
5	6	7	8	9	10	11
4	5	6	7	8	9	10
3	4	5	6	7	8	9
2	3	4	5	6	7	8
1	2	3	4	5	6	7
	1	2	3	4	5	6

X = First roll

Let X be r.v of 1^{st} roll and Y be r.v of 2^{nd} roll.

Consider $Z = X + Y$

Find $p_Z(z)$

Soln: $Z = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

$p_Z(2) = P(Z=2) = \frac{1}{36}$

$p_Z(3) = P(Z=3) = \frac{2}{36} = \frac{1}{18}$

$p_Z(4) = P(Z=4) = \frac{3}{36} = \frac{1}{12}$

$p_Z(5) = P(Z=5) = \frac{4}{36} = \frac{1}{9}$

$p_Z(6) = P(Z=6) = \frac{5}{36}$

$p_Z(7) = P(Z=7) = \frac{6}{36} = \frac{1}{6}$

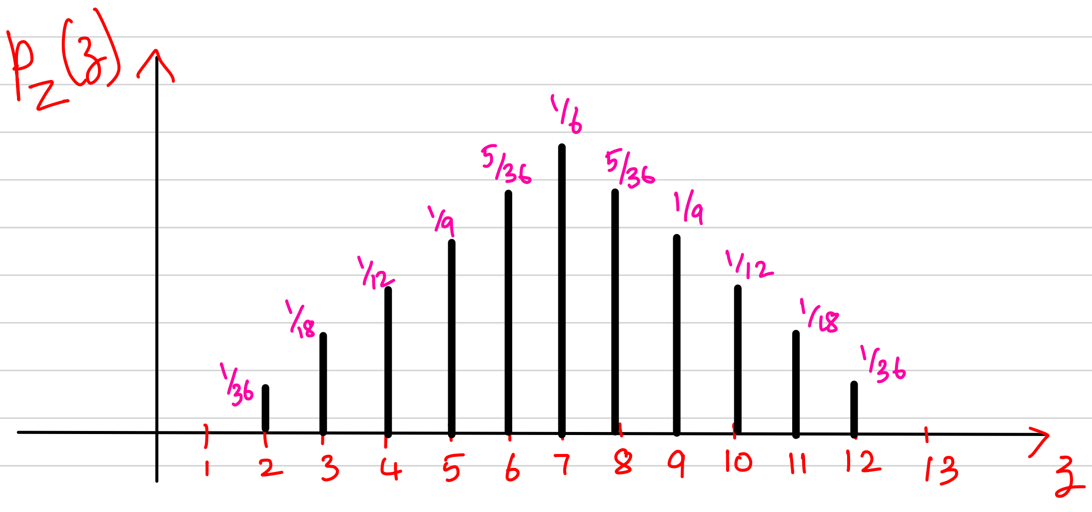
$p_Z(8) = P(Z=8) = \frac{5}{36}$

$p_Z(9) = P(Z=9) = \frac{4}{36} = \frac{1}{9}$

$p_Z(10) = P(Z=10) = \frac{3}{36} = \frac{1}{12}$

$p_Z(11) = P(Z=11) = \frac{2}{36} = \frac{1}{18}$

$p_Z(12) = P(Z=12) = \frac{1}{36}$

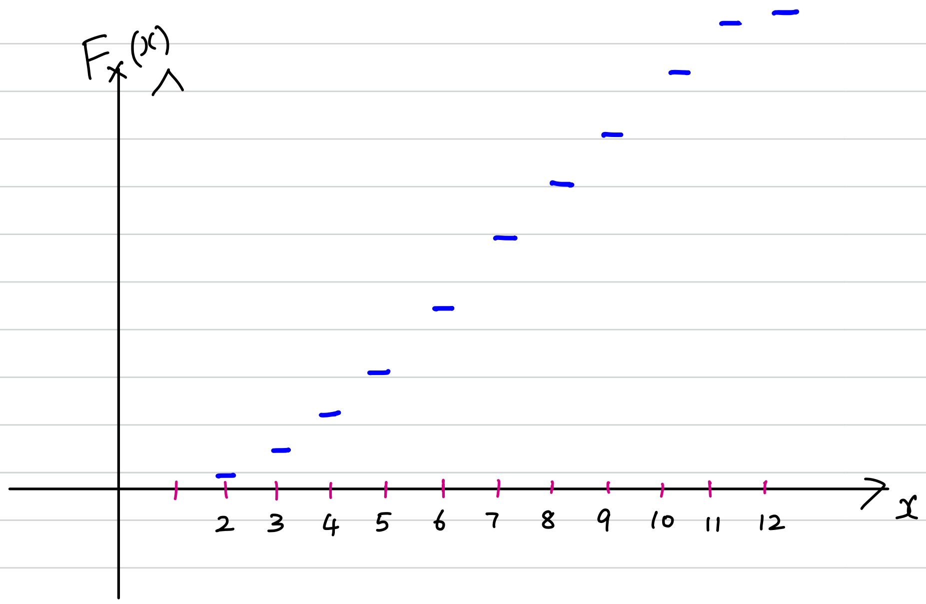


Probability distribution:

x	2	3	4	5	6	7	8	9	10	11	12
$P_X(x)$	$1/36$	$1/18$	$1/12$	$1/9$	$5/36$	$1/6$	$5/36$	$1/9$	$1/12$	$1/18$	$1/36$

Cumulative distribution

x	2	3	4	5	6	7	8	9	10	11	12
$F_X(x)$	$1/36$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}=1$



Expectation / mean of a random variable

Let X be a discrete r.v having a probability mass fn. $p_X(x)$, then the expectation, or the expected value of X, denoted by $E[X]$, is defined by

$$E[X] = \sum_x x\, p_X(x)$$

In other words, the expected value of X is a weighted average of the possible values that X can take on, each value being weighted by the probability that X assumes it.

Another motivation to the defn of Expectation is provided by the freq interpretation of probabilities.

Ex: Play a game of chance over and over (say n times).

Let X represents our winning in a single game of chance. That is, we shall win x_i units $i = 1, 2, \cdots, n$ with probabitities $p(x_i)$.

Suppose $n = 1000$ and $X = \{2, 5, 7\}$ and

win 200 times 2 units $\Rightarrow$ $P_X(2) = 0.2$
500 " 5 " $P_X(5) = 0.5$
300 " 7 units $P_X(7) = 0.7$

'Avg' gain:

$$\frac{1\cdot 200 + 5\cdot 500 + 7\cdot 300}{1000}$$

$$= 1 \times \frac{2}{10} + 5 \times \frac{5}{10} + 7 \times \frac{3}{10} = 4.8$$

Interpretation: Average in large no. of independent repetitions of the experiment.

Elementary properties of expectation

- If $X \geq 0$, then $E[X] \geq 0$
- If $a \leq X \leq b$, then $a \leq E[X] \leq b$

$$E[X] = \sum_x x\, p_X(x) \geq \sum_x a\, p_X(x) = a \sum_x p_X(x) = a$$

Illy $E[X] \leq b$.

- If c is a constant, $E[c] = c$

Expectation of a function of a random variable

Let X be a r.v. and let $Y = g(X)$.

$$E[Y] = \sum_y y\, p_Y(y)$$

$$= y_1\, p_Y(y_1) + y_2\, p_Y(y_2)$$

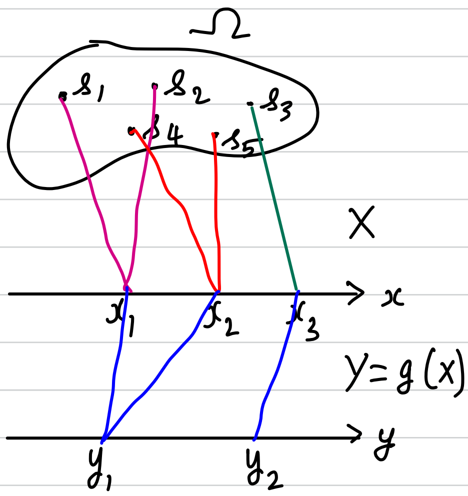
$$= y_1\left(p_X(x_1) + p_X(x_2)\right) + y_2\left(p_X(x_3)\right)$$

$$= \left[g(x_1)\, p_X(x_1) + g(x_2)\, p_X(x_2)\right] + g(x_3)\, p_X(x_3)$$

$(\because y_1 = g(x_1) = g(x_2))$

$$= \sum_{x:\, g(x) = y_1} g(x)\, p_X(x) + \sum_{x:\, g(x) = y_2} g(x)\, p_X(x)$$

$$= \sum_x g(x)\, p_X(x)$$



$g(x_1) = y_1$, $g(x_2) = y_1$

$g(x_3) = y_2$

Thus,

$$E[Y] = \sum_x g(x)\, p_X(x)$$

In particular,

If $Y = X^2$, $E[X^2] = \sum_x x^2\, p_X(x)$

Linearity of expectation :

$E[aX+b] = aE[X] + b$, where a and b are consts.

Let $g(x) = ax+b$

$$E[ax+b] = \sum_x (ax+b) P_X(x)$$

$$= a\sum_x x P_X(x) + b\sum_x P_X(x)$$

$$= aE[X] + b \qquad \left(\because \sum_x P_X(x) = 1\right)$$

In general, $E[g(X)] \neq g\, E[X]$

Variance and Standard deviation

Variance is a quantity that measures the amount of spread or the dispersion of a probability distribution (or PMF)

ie In some sense Variance quantify the amount of randomness that is present.

Definition:

If X is a random variable with mean $\mu = E[X]$, then the variance of X, denoted by $Var(X)$ is defined by

$$Var(X) = \sum_x (x-\mu)^2 P_X(x) = E[(X-\mu)^2]$$

Note: $Var(X) \geq 0$

An alternate formula for $Var(X)$:

$$Var(X) = E[X^2] - (E[X])^2$$

pf: $Var(X) = E[(X-\mu)^2]$

$= E[X^2 + \mu^2 - 2X\mu]$

$= E[X^2] + E[\mu^2] - E[2X\mu]$

$= E[X^2] + \mu^2 E[1] - 2\mu E[X]$

$= E[X^2] + \mu^2 - 2\mu^2 \qquad (\because E[1] = 1)$

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

Standard deviation: $\sigma_X = \sqrt{Var(X)}$

Properties of the Variance

$$Var(aX+b) = a^2 Var(X)$$

pf: $E[aX+b] = a E[X] + b = a\mu + b$

$\therefore Var(aX+b) = E[(aX+b-(a\mu+b))^2]$

$= E[(aX+\cancel{b}-a\mu-\cancel{b})^2]$

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

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

$= a^2 Var(X)$

Ex1: Verify if the following function can be PMF

i) $f(x) = \frac{x}{24}$ where $x = \{0, 1, 2, 3, 4\}$

ii) $f(x) = \frac{15 - x^2}{24}$ where $x = \{0, 1, 2, 3, 4\}$

Soln: $f(x)$ is a PMF if a) $f(x) \geq 0 \;\; \forall x$

b) $\sum_x f(x) = 1$

i) Clearly $f(x) \geq 0$ for all x.

$$\sum_x f(x) = \frac{0}{24} + \frac{1}{24} + \frac{2}{24} + \frac{3}{24} + \frac{4}{24} = \frac{10}{24} \neq 0$$

Thus $f(x)$ is not a PMF

ii) when $x = 4$, $f(x) = \frac{15 - 4^2}{24} = -\frac{1}{24} \ngeq 0$

$\Rightarrow$ $f(x)$ is not PMF.

Ex2: Find 'K' such that the table

x	-3	-2	-1	0	1	2	3
$P_X(x)$	K	$2K$	$3K$	$4K$	$3K$	$2K$	K

represents probability distribution. Find its expectation and variance. Also find i) $P(X \geq 0)$ ii) $P(-2 < X \leq 2)$

Soln: Given $P_X(x)$ is probability distribution. $\therefore \sum_x P_X(x) = 1$

$\Rightarrow$ $K + 2K + 3K + 4K + 3K + 2K + K = 1$

$\Rightarrow$ $16K = 1$

$\Rightarrow$ $K = 1/16$

Prob. distribution

x	-3	-2	-1	0	1	2	3
$P_X(x)$	$1/16$	$2/16$	$3/16$	$4/16$	$3/16$	$2/16$	$1/6$

Expectation, $E[X] = \sum_{x} x \, p_X(x)$

$$= -3 \times \frac{1}{16} - 2 \times \frac{2}{16} - 1 \times \frac{3}{16} + 0 \times \frac{4}{16} + 1 \times \frac{3}{16} + 2 \times \frac{2}{16} + 3 \times \frac{1}{16}$$

$$= 0$$

Variance, $Var(X) = E[X^2] - (E[X])^2$

$$= E[X^2] - 0$$

$$= \sum_{x} x^2 p_X(x)$$

$$= 9 \times \frac{1}{16} + 4 \times \frac{2}{16} + 1 \times \frac{3}{16} + 0 \times \frac{4}{16} + 1 \times \frac{3}{16} + 4 \times \frac{2}{16} + 9 \times \frac{1}{16}$$

$$= \frac{40}{16} = 2.5$$

$$P(X \geq 0) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= \frac{4}{16} + \frac{3}{16} + \frac{2}{16} + \frac{1}{16} = \frac{10}{16}$$

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

$$= \frac{3}{16} + \frac{4}{16} + \frac{3}{16} + \frac{2}{16} = \frac{12}{16}$$

Ex 3: Five defective bulbs are accidently mixed with twenty good ones. It is not possible to just look at a bulb and tell whether or not it is defective. Find the mean of the no. of defective bulbs if 4 bulbs are drawn at random from this lot.

Soln: Let X denote the no. of defective bulbs. Then

$X = \{0, 1, 2, 3, 4\}$. Since 4 bulbs are drawn at random.

No. of possible outcomes of the sample space

= No. of ways we can select 4 bulbs from 25 bulbs

$= \binom{25}{4}$ (or $^{25}C_4$)

$$P(X=0) = \frac{\text{No. of ways we can select 0 defective bulbs}}{\text{No. of ways we can select 4 bulbs}}$$

$$= \frac{\binom{20}{4}}{\binom{25}{4}} = \frac{\frac{20\times19\times18\times17}{4!}}{\frac{25\times24\times23\times22}{4!}} = \frac{969}{2530}$$

$$P(X=1) = \frac{\text{\# of way we can select 1 defective \& 3 good bulbs}}{\binom{25}{4}}$$

$$= \frac{\binom{5}{1}\times\binom{20}{3}}{\binom{25}{4}} = \frac{1140}{2530}$$

$$P(X=2) = \frac{\binom{5}{2}\times\binom{20}{2}}{\binom{25}{4}} = \frac{380}{2530}$$

$$P(X=3) = \frac{\binom{5}{3}\times\binom{20}{1}}{\binom{25}{4}} = \frac{40}{2530}$$

$$P(X=4) = \frac{\binom{5}{4}}{\binom{25}{4}} = \frac{1}{2530}$$

Prob. distribution

x	0	1	2	3	4
$P_X(x)$	$\frac{969}{2530}$	$\frac{1140}{2530}$	$\frac{380}{2530}$	$\frac{40}{2530}$	$\frac{1}{2530}$

mean, $E[X] = \sum_x x\, p_X(x)$

$= 0 \times \frac{969}{2530} + 1 \times \frac{1140}{2530} + 2 \times \frac{380}{2530} + 3 \times \frac{46}{2530} + 4 \times \frac{1}{2530}$

$= 0.8$

Ex 4: Toss a coin repeated until head occurs for the 1st time.
Let X denote the no. of tosses before head occurs.
Find i) PMF ii) $P(X)$ iii) $P(X = \text{even})$ iv) $P(X = x : 3 | x)$

Soln: Sample space $\Omega = \{H, TH, TTH, TTTH, TTTTH, \dots\}$

R.V. $X = \{1, 2, 3, 4, 5, \dots\}$

$p_X(1) = \frac{1}{2}$

$p_X(2) = \frac{1}{2^2}$

$p_X(3) = \frac{1}{2^3}$

i) In general, prob of occurence of head after n tosses

$p_X(n) = \frac{1}{2^n}, \quad n = 1, 2, 3, 4, \dots$

This is PMF.

ii) $P(X) = \sum_x p_X(x)$

$= \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \dots$

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

G.S

$a + ar + ar^2 + ar^3 + \dots$

$= \frac{a}{1-r}; \quad |ar| < 1$

iii) $P(X = \text{even}) = \sum_{x:\text{even}} p_X(x)$

$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \frac{1}{2^8} + \dots = \frac{\frac{1}{2^2}}{1 - \frac{1}{2^2}} = \frac{1}{3}$

iv) $P(X = x : 3 | x) = \sum_{3|x} p_X(x)$

$$= \frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \frac{1}{2^{12}} + \cdots$$

$$= \frac{\frac{1}{2^3}}{1 - \frac{1}{2^3}} = \frac{1}{7}$$

Ex 5: A pair of dice is thrown. Let X be the r.v. which denotes the minimum of 2 no.s which appear. Find the mean and S.D of X.

Soln: Sample space $\Omega = \{(a,b) \,/\, 1 \le a, b \le 6\}$

$X = \min(a,b) = \{1, 2, 3, 4, 5, 6\}$

# of possible outcomes, $|\Omega| = 36$

# of possible outcomes when min is 1

$= \#\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (3,1), (4,1), (5,1), (6,1)\}$

$= 11$

$\therefore P(X=1) = \frac{11}{36}$

$|||^{ly}$ $P(X=2) = \frac{9}{36}$, $P(X=3) = \frac{7}{36}$, $P(X=4) = 5/36$

$P(X=5) = 3/36$, $P(X=6) = \frac{1}{36}$

prob. distribution.

x	1	2	3	4	5	6
$p_X(x)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

mean, $E[X] = \sum_x x\, p_X(x) = 1 \times \frac{11}{36} + 2 \times \frac{9}{36} + 3 \times \frac{7}{36} + 4 \times \frac{5}{36} + 5 \times \frac{3}{36} + 6 \times \frac{1}{36}$

$$= \frac{91}{36} \approx 2.53$$

Variance, $Var(X) = E[X^2] - (E[X])^2$

$$= \sum_x x^2 p_X(x) - 2.53^2$$

$$= 1 \times \frac{11}{36} + 4 \times \frac{9}{36} + 9 \times \frac{7}{36} + 16 \times \frac{5}{36} + 25 \times \frac{3}{36} + 36 \times \frac{1}{36} - 2.53^2$$

$$\simeq 1.959$$

$$S.D = \sqrt{Var(X)} = \sqrt{1.954} \simeq 1.4$$

Ex6: If X is a discrete r.v. taking values $1, 2, 3, \ldots$ with $p_X(x) = \frac{1}{2}\left(\frac{2}{3}\right)^x$. Find $P(X = x$, being an odd no.$)$ by first establishing that $p_X(x)$ is a PMF

Soln: For $p_X(x)$ to be PMF, it has to satisfy

i) $p_X(x) \geq 0 \quad \forall x$

ii) $\sum_x p_X(x) = 1$

Clearly i) is true,

Consider $\sum_x p_X(x) = \sum_{n=1}^{\infty} \frac{1}{2}\left(\frac{2}{3}\right)^n$

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

$$= \frac{1}{2}\left(\frac{2/3}{1-2/3}\right) = 1$$

$\Rightarrow p_X(x) = \frac{1}{2}\left(\frac{2}{3}\right)^x$ is PMF.

$$P(X = x, \text{ an odd no.}) = \sum_{2 \nmid n} \frac{1}{2}\left(\frac{2}{3}\right)^n = \frac{1}{2}\left[\frac{2}{3} + \left(\frac{2}{3}\right)^3 + \left(\frac{2}{3}\right)^5 + \cdots\right]$$

$$= \frac{1}{2}\left(\frac{2/3}{1-(2/3)^2}\right) = 0.6$$

Ex 7: Determine probability mass function for the following cumulative distribution function.

$$F_X(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \le x < 0 \\ 0.7 & 0 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

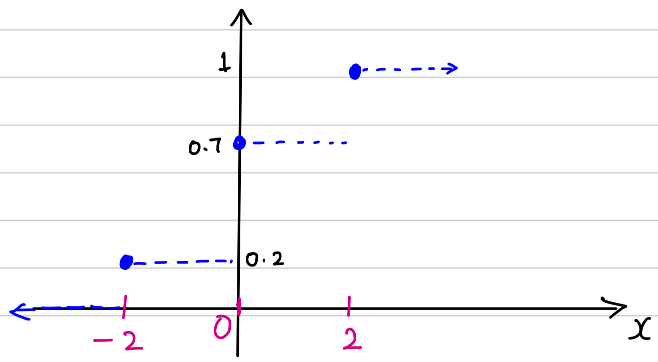
Soln: Clearly, $X = \{-2, 0, 2\}$

$F_X(-2) = 0.2$

$F_X(0) = 0.7$

$F_X(2) = 1$

Let us plot $F_X(x)$.



pmf at each pt. is the change in the cumulative distribution function at the point.

$P_X(-2) = 0.2 - 0 = 0.2$

$P_X(0) = 0.7 - 0.2 = 0.5$

$P_X(2) = 1 - 0.7 = 0.3$

Continuous random variables

There exist random variables whose set of possible values is uncountable.

Ex. i) The time that a train arrives at a specified stop.

ii) Lifetime of a transistor

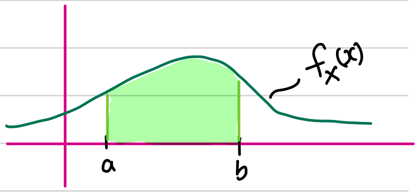
iii) The time taken for the milk to become curd.

Let X be such a random variable. We say that X is a continuous r.v. if there exists a nonnegative function f_X, defined for all $x \in (-\infty, \infty)$, having the following properties

1) $f_X(x) \geq 0, \ \forall x \in \mathbb{R}$

2) $\int_{-\infty}^{\infty} f_X(x)\,dx = P(X \in (-\infty, \infty)) = 1$

3) $P(a \leq X \leq b) = \int_a^b f_X(x)\,dx$



The function is called the probability density function of the r.v. X. ($f_X(x)$ tell us the probability per unit length)

If we let $a=b$ in (3), we get

$$P(X=a) = \int_a^a f_X(x)\,dx = 0$$

In other words, the equation states that the probability of a continuous r.v at any fixed value is zero.

Hence, $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$

Cumulative distribution function or distribution function

Let X be a continuous r.v. and let $f_X(x)$ be the PDF. Then the cumulative distribution function CDF is defined by

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

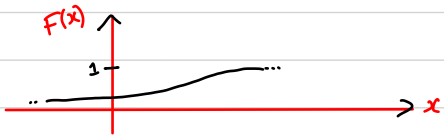
From above defn it is clear that

i) $P(a \le X \le b) = F_X(b) - F_X(a)$

ii) By Fundamental thm of integral calculus

$$f_X(x) = \frac{d}{dx} F_X(x).$$

iii) If $a \le b$, then $F_X(a) \le F_X(b)$, that is $F_X(x)$ is an increasing fn.



Expectation and Variance of Continuous random variables

If X is a continuous r.v. having probability density function $f_X(x)$, then the expected value of X is given by

$$E[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

Properties:

- If $X \ge 0$, then $E[X] \ge 0$
- If $a \le X \le b$, then $a \le E[X] \le b$
- If c is a constant, $E[c] = c$

If X is a continuous r.v. with PDF $f_X(x)$, the for any real valued function, g

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$$

- If a and b are constants, then

$$E[aX+b] = a\,E[X] + b$$

If X is continuous r.v. with expected value μ, then Variance of X is

$$\mathrm{Var}(X) = E[(X-\mu)^2]$$

The alternate formula,

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

- If a and b are constants, then

$$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$$

Ex 1: Suppose that X is a continuous r.v whose probability density function is given by

$$f_X(x) = \begin{cases} c(4x-2x^2) & 0<x<2 \\ 0 & \text{otherwise} \end{cases}$$

a) What is the value of C?
b) Find $P(X>1)$

Soln: If $f_X(x)$ is PDF, then i) $f_X(x) \geq 0$ and

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

a) From ii), $\int_{-\infty}^{\infty} c(4x-2x^2)\,dx = 1$

$$\Rightarrow \int_{-\infty}^{0} 0\,dx + \int_{0}^{2} c(4x-2x^2)\,dx + \int_{1}^{\infty} 0\,dx = 1$$

$$\Rightarrow \int_{0}^{2} c(4x-2x^2)\,dx = 1$$

$$\Rightarrow c\left[2x^2 - \frac{2}{3}x^3\right]_0^2 = 1 \Rightarrow c\left(\frac{8}{3}\right) = 1 \Rightarrow c = \frac{3}{8}$$

b) $P(X \geq 1) = \int_1^\infty f_X(x)\,dx = \int_1^\infty c\,(4x - 2x^2)\,dx$

$$= \int_1^2 \frac{3}{8}(4x - 2x^2)\,dx$$

$$= \frac{3}{8}\left[2x^2 - \frac{2}{3}x^3\right]_1^2 = \frac{3}{8}\left[\frac{8}{3} - \frac{4}{3}\right]$$

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

Ex2: The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that

(a) a computer will function between 50 and 150 hours before breaking down?
(b) it will function for fewer than100 hours?

Soln: Since $f(x)$ is PDF

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

$$= \int_0^\infty \lambda e^{-x/100}\,dx = 1$$

$$\Rightarrow \lambda \left[\frac{e^{-x/100}}{-1/100}\right]_0^\infty = 1$$

$$\Rightarrow \lambda \left[0 + \frac{1}{1/100}\right] = 0$$

$$\left[\frac{e^{-x/100}}{-1/100}\right]_0^\infty = \lim_{t \to \infty} \left[\frac{e^{-x/100}}{-1/100}\right]_0^t$$

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

a) $P(50 \leq X \leq 150) = \int_{50}^{150} f(x)\,dx = \int_{50}^{150} \frac{1}{100} e^{-x/100}\,dx$

$$= \left[-e^{-x/100}\right]_{50}^{150} \approx 0.383$$

b) $P(X \le 100) = \int_{-\infty}^{100} f(x)\,dx = \int_{0}^{100} \frac{1}{100} e^{-x/100}\,dx$

$$= -e^{-x/100}\Big]_0^{100}$$

$$= -e^{-1} + 1 \approx 0.632$$

Ex 3 : Find the mean and Variance of the PDF $f_X(x) = \frac{e^{-|x|}}{2}$

soln : $f_X(x) = \begin{cases} \frac{e^{-x}}{2}, & x \ge 0 \\ \frac{e^{x}}{2}, & x < 0 \end{cases}$

mean, $E[X] = \int_{-\infty}^{\infty} x\, f_X(x)\,dx = \frac{1}{2}\int_{-\infty}^{0} x e^{x}\,dx + \frac{1}{2}\int_{0}^{\infty} x e^{-x}\,dx$

$$= xe^{x} - e^{x}\Big]_{-\infty}^{0} + xe^{-x} + e^{-x}\Big]_0^{\infty}$$

$$= -1 + 1$$

$$= 0$$

Variance, $Var(x) = E[x^2] - \cancel{E[x]^2}\ 0$

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

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

$$= \frac{1}{2}\left\{ x^2e^{x} - 2xe^{x} + 2e^{x}\Big]_{-\infty}^{0} + x^2e^{-x} + 2xe^{-x} - 2e^{-x}\Big]_0^{\infty} \right\}$$

$$= \frac{1}{2}\{2 + 2\} = \frac{4}{2} = 2$$

Ex 4 : In a certain city, the daily consumption of electric power (in million kW/hr) is a continous r.v. having the PDF.

$$f_X(x) = \begin{cases} \frac{1}{9} x e^{-x/3} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

If the cities power plant has a daily capacity of 12 million kW/hr. What is the probability that this power supply will be insufficient on any given day.

Soln: Given PDF, $f_X(x) = \begin{cases} \frac{1}{9} x e^{-x/3} & x \geq 0 \\ 0 & x < 0 \end{cases}$

The power supply will be insufficient on any given day if power consumption is > 12 million kW/hr.

and $P(X > 12) = \int_{12}^{\infty} f_X(x)\, dx$

$$= \int_{12}^{\infty} \frac{1}{9} x e^{-x/3}\, dx$$

$$= \frac{1}{9} \left[x \frac{e^{-x/3}}{-1/3} - \frac{e^{-x/3}}{1/9} \right]_{12}^{\infty}$$

$$= \frac{1}{9} \left[+36 e^{-4} + 9 e^{-4} \right] = \frac{45}{9} e^{-4}.$$

Ex 5: The length of time (in minutes) that a certain lady speaks on telephone is found to be a continous r.v. with PDF

$$f_X(x) = \begin{cases} A e^{-x/5} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find (i) A. Find the probability that she will speak on the phone (ii) more than 10 min (iii) less than 5 min iv) between 5 and 10 min

Soln: i) Given PDF, $f_X(x) = \begin{cases} A e^{-x/5}, & x \geq 0 \\ 0 & x < 0 \end{cases}$

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

$$\Rightarrow \int_{0}^{\infty} A e^{-x/5}\, dx = 1$$

$$\Rightarrow A \left[\frac{e^{-x/5}}{-1/5} \right]_{0}^{\infty} = 1$$

$$\Rightarrow A = \frac{1}{5}$$

ii) $P(X > 10) = \int_{10}^{\infty} f_X(x)\, dx = \int_{10}^{\infty} \frac{1}{5} e^{-x/5}\, dx = e^{-2} = 0.1353$

iii) $P(X<5) = \int_{-\infty}^{5} f_X(x)\,dx = \int_{0}^{5} \frac{1}{5} e^{-x/5}\,dx = -e^{-1}+1$

$= 0.6321$

iv) $P(5 \le X \le 10) = \int_{5}^{10} f_X(x)\,dx = \int_{5}^{10} \frac{1}{5} e^{-x/5}\,dx = -e^{-2}+e^{-1}$

$= 0.2325$

Ex6 : The PDF of a continuous r.v. is

$$f_X(x) = \begin{cases} k\,x(1-x)\,e^{x} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find k and evaluate mean and S.D. of the distribution.

Soln: Given $f_X(x)$ is PDF

$\therefore \int_{-\infty}^{\infty} f_X(x)\,dx = 1 \Rightarrow \int_{0}^{1} k\,x(1-x)\,e^{x}\,dx = 1$

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

$\Rightarrow k\left[+e - 2e + 1 + 2\right] = 1$

$\Rightarrow k = \frac{1}{3-e}$

mean, $E[X] = \int_{0}^{1} \frac{1}{3-e}\, x^2(1-x)\,e^{x}\,dx$

$= \frac{1}{3-e}\left[x^2(1-x)e^{x} - (2x-3x^2)e^{x} + (2-6x)e^{x} + 6e^{x}\right]_0^1$

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

$= \frac{3e-8}{3-e} \approx 0.5496$

$Var(X) = E[X^2] - E[X]^2 = \cdots$

Ex7 : The density function of X is given by

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Let $Y = e^X$. i) Find PDF of Y ii) $E(Y)$

Soln: Cumulative distribution function of Y is

$$F_Y(y) = P(Y \le y) = P(e^X \le y)$$

$$= P(X \le \log y)$$

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

$$= \cancel{\int_{-\infty}^{0} f_X(x)\,dx} + \int_{0}^{\log y} f_X(x)\,dx$$

when $0 \le x \le 1$, $1 \le y \le e \Rightarrow 0 \le \log y \le 1$

$$\Rightarrow F_Y(y) = \int_0^{\log y} 1\,dx = \log y$$

By differentiating, we get

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{y}, \qquad 1 \le y \le e$$

Hence, $$E[Y] = \int_1^e y\, f_Y(y)\,dy = e - 1.$$

or

$$E[Y] = \int_0^1 e^x\,dx = e - 1.$$

Ex8 : The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \le 100 \\ \frac{100}{x^2} & x > 100 \end{cases}$$

i) What is the cumulative distribution function of X?

ii) What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation?

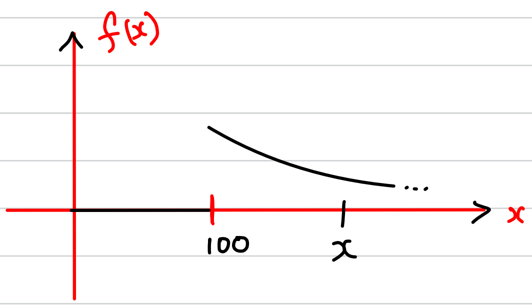
(hint : Assume that the events E_i ,i = 1,2,3,4,5, that the ith such tube will have to be replaced within this time are independent.)

Soln: Given PDF

$$f(x) = \begin{cases} 0 & x \le 100, \\ \frac{100}{x^2} & x > 100. \end{cases}$$

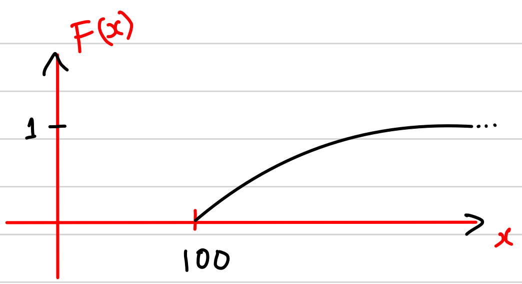
i) Cumulative distribution

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



$$= \int_{100}^{x} f(x)\,dx$$

$$= \int_{100}^{x} \frac{100}{x^2}\,dx$$



$$= 100 \left. \frac{-1}{x} \right]_{100}^{x}$$

$$= 100 \left(\frac{-1}{x} + \frac{1}{100} \right)$$

$$F(x) = \frac{x-100}{x}, \quad x \geqslant 100$$

ii) From the above st. mt., we have

$$P(E_i) = P(X \le 150)$$

$$= F(150) = \frac{150-100}{150} = \frac{1}{3}$$

From independence of the events E_i, the desired prob. is

$$\binom{5}{2} \cdot \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 = \frac{80}{243} = 0.3292$$

Jointly distributed random variables

Joint Discrete random variables, Joint Probability mass function

Consider a probabilistic model with two discrete r.vs. X and Y with probability mass functions p_X and p_Y, respectively.

For any $x \in X$ and $y \in Y$, the probability of occurence of x and y is defined by

$$p_{X,Y}(x,y) = P(X=x \text{ and } Y=y)$$

Usually $P(X=x \text{ and } Y=y)$ is written as $P(X=x, Y=y)$

is called Joint Probability mass function (JPMF).

It satisfies,

i) $p_{X,Y}(x,y) \geq 0$ for all x, y

ii) $\sum_x \sum_y p_{X,Y}(x,y) = 1$

Suppose $X = \{x_1, x_2, x_3, \dots\}$, $Y = \{y_1, y_2, y_3, \dots\}$.

The $p_{X,Y}(x,y)$ can be displayed in the form of the foll. table

$p_{X,Y}(x,y)$		Y: y_1	y_2	y_3	$\dots$
X	x_1	p_{11}	p_{12}	p_{13}	$\dots$
	x_2	p_{21}	p_{22}	p_{23}	$\dots$
	x_3	p_{31}	p_{32}	p_{33}	$\dots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Notation : $p_{X,Y}(x_i, y_j) = p_{ij}$

Once we have joint PMF $p_{X,Y}(x,y)$, we can find individual PMF p_X and p_Y, they are called marginal PMFs of X and Y, respectively. From the table, for any i and j

$$p_X(x_i) = p_{i1} + p_{i2} + p_{i3} + \cdots = \sum_y p_{X,Y}(x_i, y)$$

$$p_Y(y_j) = p_{1j} + p_{2j} + p_{3j} + \cdots = \sum_x p_{X,Y}(x, y_j)$$

Note: If X and Y are independent, then

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

More than two random variables. Let X, Y, Z be r.vs.

Then

- $p_{X,Y,Z}(x,y,z) = P(X=x \text{ and } Y=y \text{ and } Z=z)$
- $\sum_x \sum_y \sum_z p_{X,Y,Z}(x,y,z) = 1$
- $p_{X,Y}(x,y) = \sum_z p_{X,Y,Z}(x,y,z)$
- $p_{Y,Z}(y,z) = \sum_x p_{X,Y,Z}(x,y,z)$
- $p_{Z,X}(z,x) = \sum_y p_{X,Y,Z}(x,y,z)$

Functions of multiple r.vs.

$Z = g(X,Y)$

PMF: $p_Z(z) = P(Z=z) = P(g(X,Y)=z)$

$$= \sum_{(x,y):\, g(x,y)=z} p_{X,Y}(x,y)$$

Expected value rule:

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y)$$

Linearity of expectation

$$E[X+Y] = E[X] + E[Y]$$

$$E[X+Y] = \sum_x \sum_y (x+y)\, p_{X,Y}(x,y)$$

$$= \sum_x \sum_y x\, p_{X,Y}(x,y) + \sum_x \sum_y y\, p_{X,Y}(x,y)$$

$$= \sum_x x \sum_y p_{X,Y}(x,y) + \sum_y y \sum_x p_{X,Y}(x,y)$$

$$= \sum_x x\, p_X(x) + \sum_y y\, p_Y(x)$$

$$= E[X] + E[Y]$$

In general, $E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$

Covariance

Covariance of X and Y is denoted by cov(X, Y), is defined by

$$\text{Cov}(X,Y) = \sum_x \sum_y (x-\mu_X)(y-\mu_Y)\, p_{X,Y}(x,y)$$

$$= E[(X-\mu_X)(Y-\mu_Y)]$$

where $\mu_X = E[X]$ and $\mu_Y = E[Y]$.

Note: 1) If $Y = X$, then

$$\text{Cov}(X,X) = E[(X-\mu_X)^2] = \text{Var}(X)$$

2) $\text{Cov}(X,Y) = E[XY] - E[X]\,E[Y]$

From the defn $\text{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

$$= E[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y]$$

$$= E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X\mu_Y$$

$$= E[XY] - \mu_Y\mu_X - \cancel{\mu_X\mu_Y} + \cancel{\mu_X\mu_Y}$$

$$= E[XY] - E[X]\,E[Y]$$

Correlation Coefficient of X and Y

For r.vs X and Y, the correlation coefficient is

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X\, \sigma_Y}$$

where $\sigma_X = \sqrt{\text{Var}(X)}$, SD of X

$\sigma_Y = \sqrt{\text{Var}(Y)}$, SD of Y

For any two r.vs X and Y,

$$-1 \le \rho(X,Y) \le +1$$

Ex 1: A joint distribution of two r.vs X and Y is given by the following table:

$X \backslash Y$	-4	2	7
1	$1/8$	$1/4$	$1/8$
5	$1/4$	$1/8$	$1/8$

Are X and Y independent? Determine (i) marginal distribution of X and Y. (ii) $E[X]$ and $E[Y]$, (iii) $E[XY]$ (iv) $Cov(X,Y)$ and $\rho(X,Y)$

Soln. i) Marginal distribution of X and Y are:

x	1	5
p_X	$1/2$	$1/2$

y	-4	2	7
p_Y	$3/8$	$3/8$	$2/8$

X and Y are independent if $p_{X,Y}(x,y) = p_X(x)\, p_Y(y) \;\forall x,y$

But $p_{X,Y}(1,-4) = 1/8$ and $p_X(1)\, p_Y(-4) = 1/2 \times 3/8$

$\Rightarrow p_{X,Y}(1,-4) \neq p_X(1)\, p_Y(-4)$

$\therefore$ X and Y are not independent.

ii) $E[X] = \sum_x x\, p_X(x) = 1 \times \frac{1}{2} + 5 \times \frac{1}{2} = 3$

$E[Y] = \sum_y y\, p_Y(y) = -4 \times 3/8 + 2 \times \frac{3}{8} + 7 \times \frac{2}{8} = 1$

iii) $E[XY] = \sum_x \sum_y xy\, p_{XY}(x,y)$

$= 1 \times -4 \times \frac{1}{8} + 1 \times 2 \times \frac{1}{4} + 1 \times 7 \times \frac{1}{8} + 5 \times -4 \times \frac{1}{4} + 5 \times 2 \times \frac{1}{8}$

$+ 5 \times 7 \times \frac{1}{8} = \frac{3}{2}$

iii) $Cov(X,Y) = E[XY] - E[X]\,E[Y]$

$= 3/2 - 3 \times 1 = -3/2$

$Var(X) = E[X^2] - (E[X])^2 = 1^2 \times \frac{1}{2} + 5^2 \times \frac{1}{2} - 3^2 = 4$

$Var(Y) = E[Y^2] - (E[Y])^2 = (-4)^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 7^2 \times \frac{2}{8} - 1^2 = 18.75$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{-3/2}{\sqrt{4} \times \sqrt{18.75}} = -0.1732$$

Ex 2: The joint probability fn for two discrete r.vs X and Y is given by

$$p(x,y) = \begin{cases} c(2x+y) & \text{if } x, y \in \mathbb{Z} \text{ and } 0 \le x \le 2, \ 0 \le y \le 3, \\ 0 & \text{otherwise} \end{cases}$$

Find (i) the value of the constant c
(ii) $P(X=2, Y=1)$ (iii) $P(X \ge 1, Y \le 2)$ (iv) $P(X+Y \le 1)$
(v) $P(X+Y > 1)$

Soln: (i) If $p(x,y)$ is joint PMF, then $\sum_x \sum_y p(x,y) = 1$

X \ Y	0	1	2	3
0	0	c	2c	3c
1	2c	3c	4c	5c
2	4c	5c	6c	7c

$$\sum_x \sum_y p(x,y) = 0 + c + 2c + 3c + 2c + 3c + 4c + 5c + 4c + 5c + 6c + 7c = 1$$

$$\Rightarrow 42c = 1 \Rightarrow c = \frac{1}{42}$$

ii) $P(X=2, Y=1) = 5 \times \frac{1}{42}$

ii) $P(X \ge 1, Y \le 2) = (2+3+4+4+5+6)\frac{1}{42} = \frac{24}{42} = 4/7$

iii) $P(X+Y \le 1) = (0+1+2+3)\frac{1}{42} = \frac{6}{42} = \frac{1}{7}$

iv) $P(X+Y > 1) = 1 - P(X+Y \le 1) = 1 - \frac{1}{7} = \frac{6}{7}$

Ex3: A coin is tossed three times. Let X be equal to 0 or 1 according as a head or a tail occurs on the first toss.
Let Y be equal to the total no. of heads which occur

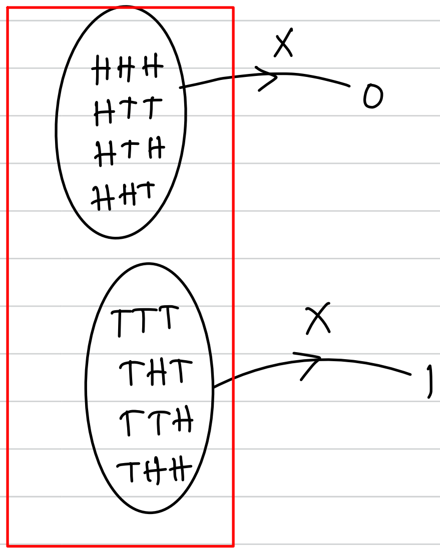
Determine (i) The marginal distribution of X and Y
(ii) The joint distribution of X and Y
(iii) Expected values of X, Y, X+Y and XY
(iv) Cov(X,Y) and $\rho(X,Y)$

Soln: Sample space $\Omega = \{HHH, HTT, HTH, HHT, THH, THT, TTH, TTT\}$

$X = 0$ if head occurs on the 1st toss
$X = 1$ if tail occurs on the 1st toss

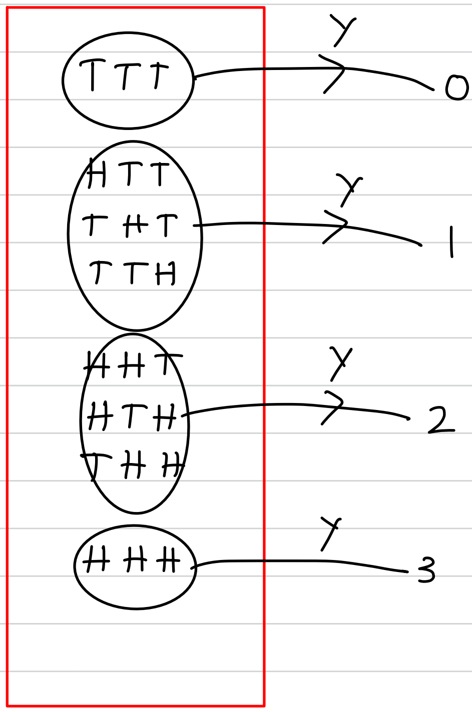
$\therefore X = \{0, 1\}$

$Y = \{0, 1, 2, 3\}$



ii) Joint PMF

X \ Y	0	1	2	3
0	0	1/8	2/8	1/8
1	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0



i) Marginal distributions

x	0	1
$P_X(x)$	1/2	1/2

y	0	1	2	3
$P_Y(y)$	1/8	3/8	3/8	1/8

iii) $E[X] = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = 1/2$

$E[Y] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = 1.5$

$$E[X+Y] = E[X] + E[Y] = \tfrac{1}{2} + \tfrac{3}{2} = 2$$

$$E[XY] = 0\times0\times0 + 0\times1\times\frac{1}{8} + 0\times2\times\frac{2}{8} + 0\times3\times\frac{1}{8}$$

$$+ 1\times0\times\frac{1}{8} + 1\times1+\frac{2}{8} + 1\times2\times\frac{1}{8} + 1\times3\times0$$

$$= \tfrac{1}{2}$$

iv) $$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y]$$

$$= \tfrac{1}{2} - \tfrac{1}{2}\times\tfrac{3}{2} = -\tfrac{1}{4}$$

$$\mathrm{Var}(X) = 0^2\times\frac{1}{2} + 1^2\times\frac{1}{2} - \left(\frac{1}{2}\right)^2 = \tfrac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$\mathrm{Var}(Y) = 0^2\times\frac{1}{8} + 1^2\times\frac{3}{8} + 2^2\times\frac{3}{8} + 3^2\times\frac{1}{8} - \left(\frac{3}{2}\right)^2 = \frac{3+12+9}{8} - \frac{9}{4} = 3 - \frac{9}{4} = \frac{3}{4}$$

$$\therefore\ \sigma_X = \sqrt{\mathrm{Var}(X)} = \tfrac{1}{2}$$

$$\sigma_Y = \sqrt{\mathrm{Var}(Y)} = \frac{\sqrt{3}}{2}$$

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\ \sigma_Y} = \frac{-\tfrac{1}{4}}{\tfrac{1}{2}\,\tfrac{\sqrt{3}}{2}} = -\frac{1}{\sqrt{3}}$$

Ex4 · Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let Xi equal 1 if the ith ball selected is white, and let it equal 0 otherwise. Give the joint probability mass function of

a) X_1, X_2 ; b) X_1, X_2, X_3

Soln: The r.v. X_i is defined as

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ ball is white} \\ 0 & \text{otherwise} \end{cases}$$

$\Rightarrow \quad X_i = \{0, 1\}, \quad i = 1, 2, 3$

$P(X_1=0 \text{ and } X_2=0) = P(X_1=0) \cdot P(X_2=0) \text{ given } X_1=0$

$= $ Prob that 1^{st} ball is red $\times$ Prob that 2^{nd} ball is red given 1^{st} ball is red

$$= \frac{8}{13} \times \frac{7}{12}$$

lllle $P(X_1=0 \text{ and } X_2=1) = \dfrac{8\times5}{13\times12} = \dfrac{10}{39}$

$P(X_1=1 \text{ and } X_2=0) = \dfrac{5\times8}{13\times12} = \dfrac{10}{39}$

$P(X_1=1 \text{ and } X_2=1) = \dfrac{5\times4}{13\times12} = \dfrac{5}{39}$

i.e

$X_1 \backslash X_2$	0	1
0	$\frac{14}{39}$	$\frac{10}{39}$
1	$\frac{10}{39}$	$\frac{5}{39}$

Marginal distribution of X_1

$P(X_1=0) = \dfrac{24}{39}, \quad P(X_1=1) = \dfrac{15}{39}$

Marginal distributn of X_2

$P(X_2=0) = \dfrac{24}{39}, \quad P(X_2=1) = \dfrac{15}{39}$

b) $X_1 = \{0, 1\}$, $X_2 = \{0, 1\}$, $X_3 = \{0, 1\}$

$\therefore P(X_1 = 0, X_2 = 0, X_3 = 0) =$ Prob. that 1st ball is red
× Prob that 2nd ball is red given 1st ball is red
× prob that 3rd ball is red given 1st and 2nd ball is red

$= \frac{8}{13} \times \frac{7}{12} \times \frac{6}{11}$

Joint prob. distribution is

X_1	X_2	X_3	$P_{X,Y}(x_1, x_2, x_3)$
0	0	0	$\frac{8}{13} \times \frac{7}{12} \times \frac{6}{12}$
0	0	1	$\frac{8}{13} \times \frac{7}{12} \times \frac{5}{11}$
0	1	0	$\frac{8}{13} \times \frac{5}{12} \times \frac{7}{11}$
0	1	1	$\frac{8}{13} \times \frac{5}{12} \times \frac{4}{11}$
1	0	0	$\frac{5}{13} \times \frac{8}{12} \times \frac{7}{12}$
1	0	1	$\frac{5}{13} \times \frac{8}{12} \times \frac{4}{11}$
1	1	0	$\frac{5}{13} \times \frac{4}{12} \times \frac{8}{11}$
1	1	1	$\frac{5}{13} \times \frac{4}{12} \times \frac{3}{11}$

Conditional PMFs

Let X and Y be two discrete r.v.s. with joint probability function $p_{X,Y}(x,y)$.

The conditional probability mass function of Y given $X=x$

$$p_{Y|X}(y|x) = P(Y=y \mid X=x)$$

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

$$= \frac{p_{X,Y}(x,y)}{p_X(x)} \quad \text{defined for } x \text{ such that } p_X(x)>0$$

Multiplicative rule:

$$P(Y=y,\ X=x) = P(X=x)\cdot P(Y=y \mid X=x)$$
$$= P(Y=y)\cdot P(X=x \mid Y=y)$$

In joint PDF notation:

$$p_{X,Y}(x,y) = p_X(x)\cdot p_{Y|X}(y|x) = p_Y(y)\cdot p_{X|Y}(x|y)$$

Ex: Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find the conditional distribution of X, given that $Y=1$, and use it to determine $P(X=0 \mid Y=1)$

Soln: The possible pairs of values (x,y) are

$$\{(0,0),(0,1),(0,2),(1,0),(1,1),(2,0)\}$$

Joint prob. distribution

$$p_{X,Y}(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}}$$

$$= \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{28}$$

$p_{X,Y}(x,y)$		$X=0$	$X=1$	$X=2$
Y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0
	2	$\frac{1}{28}$	0	0

$P(Y=1) = \frac{3}{7}$

$$p_{X|Y}(x|1) = P(X=x \mid Y=1) = \frac{P(X=x, Y=1)}{P(Y=1)} = p_{X,Y}(x,1) \times \frac{7}{3}$$

x	0	1	2
$p_{X\|Y}(x\|1)$	$\frac{3}{14} \times \frac{7}{3} = \frac{1}{2}$	$\frac{3}{4} \times \frac{7}{3} = \frac{1}{2}$	$0 \times \frac{7}{3} = 0$

Finally, $P(X=0 \mid Y=1) = 1/2$

Properties of Conditional PMFs

1) $p_{X|Y}(x|y) \geq 0$

2) $\sum_x p_{X|Y}(x|y) = 1$

Conditional expectation and Conditional Variance

Conditional expectation of X given $Y=y$ is

$$E[X \mid Y=y] = \sum_x x\, p_{X|Y}(x|y)$$

If $Z = g(X)$, then

$$E[Z \mid Y=y] = \sum_x g(x)\, p_{X|Y}(x|y)$$

Conditional Variance of X given $Y=y$ is

$$\mathrm{Var}(X \mid Y=y) = E[X^2 \mid Y=y] - (E[X \mid Y=y])^2$$

Total Probability and expectation theorem

$$p_X(x) = \sum_y p_{X,Y}(x,y) = \sum_y p_Y(y)\, p_{X|Y}(x|y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y) = \sum_x p_X(x)\, p_{Y|X}(y|x)$$

$E[X] = \sum_y p_Y(y)\ E[X \mid Y=y]$ $\qquad$ $E[Y] = \sum_x p_X(x)\ E[Y \mid X=x]$

Independence

For r.vs X and Y, the following properties are equivalent, and X and Y are independent.

1) $p_{XY}(x,y) = p_X(x) \cdot p_Y(y)$ for all x and y
2) $p_{X|Y}(x|y) = p_X(x)$ for all x with $p_Y(y) \neq 0$
3) $p_{Y|X}(y|x) = p_Y(y)$ for all y with $p_X(x) \neq 0$

Independence and expectation

If r.vs X and Y are independent, then

$$E[XY] = E[X] \cdot E[Y]$$

$g(X)$ and $h(Y)$ are also independent:

$$E[g(X) \cdot h(Y)] = E[g(X)] \cdot E[h(Y)].$$

pf: $E[XY] = \sum_x \sum_y xy\ p_{X,Y}(x,y)$

$= \sum_x \sum_y xy\ p_X(x) \cdot p_Y(y)$ $\quad$ ($\because$ X and Y are independent)

$= \sum_x x\ p_X(x) \cdot \sum_y y\ p_Y(y)$

$= E[X] \cdot E[Y]$

Note: In general, $E[XY] \neq E[X]\,E[Y]$

Independence and Variance

If X and Y are independent, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

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

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

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

$= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2$ (∵ X and Y are independent)

$= Var(X) + Var(Y)$

Ex: Let X denote the number of times a certain numerical control machine will malfunction: 1, 2, or 3 times on any given day. Let Y denote the number of times a technician is called on an emergency call. Their joint probability distribution is given as

$P_{X,Y}(x,y)$		x = 1	x = 2	x = 3
y	1	0.05	0.05	0.10
	3	0.05	0.10	0.35
	5	0.00	0.2	0.10

i) Find $P(Y=3 \mid X=2)$ ii) $E[X \mid Y=3]$ iii) $Var[X \mid Y=3]$

Soln: i) $P(Y=3 \mid X=2) = \dfrac{P(Y=3, X=2)}{P(X=2)} = \dfrac{0.10}{0.05+0.10+0.2} = \dfrac{2}{7}$

ii) $P(Y=3) = 0.05 + 0.1 + 0.35 = 0.5$

$P(X=1 \mid Y=3) = \dfrac{0.05}{0.5} = \dfrac{1}{10}$, $P(X=2 \mid Y=3) = \dfrac{0.1}{0.5} = \dfrac{1}{5}$

$P(X=3 \mid Y=3) = \dfrac{0.35}{0.5} = \dfrac{7}{10}$

x	1	2	3
$P_{X\|Y}(x\|3)$	1/10	1/5	7/10

$E[X \mid Y=3] = \sum_x x\, P_{X|Y}(x|3) = 1 \times \frac{1}{10} + 2 \times \frac{1}{5} + 3 \times \frac{7}{10} = 2.6$

$E[X^2 \mid Y=3] = \sum_x x^2 P_{X|Y}(x|3) = 1^2 \times \frac{1}{10} + 2^2 \times \frac{1}{5} + 3^2 \times \frac{7}{10} = 7.2$

$Var[X \mid Y=3] = E[X^2 \mid Y=3] - (E[X \mid Y=3])^2 = 7.2 - 6.76 = 0.44$

Joint continuous random variables, Joint probability density fn.

Let X and Y be continuous r.vs with prob. density fns f_X defined on $x \in \mathbb{R}$ and f_Y defined on $y \in \mathbb{R}$.

This two random variables are jointly continuous if they can be described by a joint prob. density function, $f_{X,Y}(x,y)$ defined over 2-D plane, satisfying the following properties.

i) $f_{X,Y}(x,y) \geq 0$ for all x, y

ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1$

iii) For any region R of two-dimensional space,

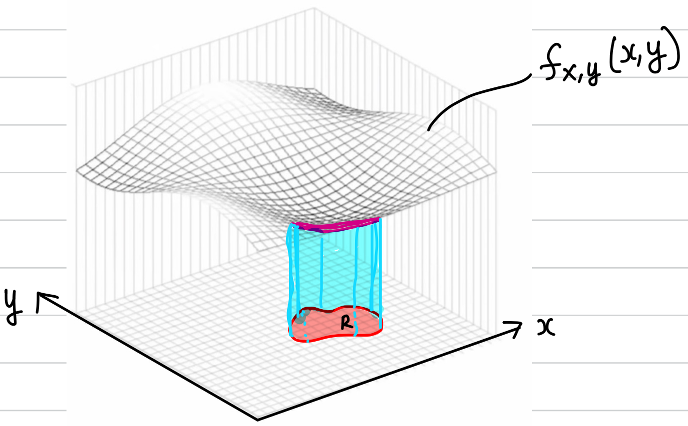
$$P((x,y) \in R) = \iint_R f_{X,Y}(x,y)\, dA$$

Visualizing a joint PDF, $f_{X,Y}(x,y)$

Prob of $(x,y) \in R$ is the volume of a solid bounded below the surface $f_{X,Y}(x,y)$ and above the region R.

Thus, $f_{X,Y}(x,y)$ tell us the probability per unit area.

Note:
If area$(R) = 0$, then $P((X,Y) \in R) = 0$



Plot of $f_{X,Y}(x,y)$

i.e prob. at a point is zero

prob. on a line is zero (This means X and Y are not jointly continuous)

Marginal probability density function

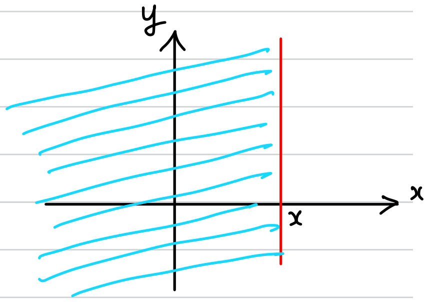
From joint PDF, $f_{X,Y}(x,y)$ we can find PDFs of individual r.vs

$f_X(x)$ and $f_Y(y)$ called marginal PDF.

$$f_X(x) = \int_{y=-\infty}^{\infty} f_{X,Y}(x,y)\, dy \quad \text{and} \quad f_Y(y) = \int_{x=-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$

pf: To find $f_X(x)$, first let us find cumulative density fn $F_X(x)$.

WKT $F_X(x) = P(X \leq x)$



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

$$f_X(x) = \frac{d}{dx} F_X(x)$$

$$= \frac{d}{dx} \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy \right] dx$$

$$= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$

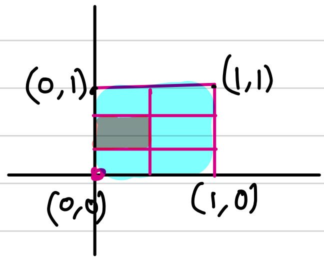
Ex : A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y , respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

a) Find $P((X,Y) \in A)$, where $A = \{(x,y) \mid 0 < x < \frac{1}{2},\ \frac{1}{4} < y < \frac{1}{2}\}$

b) Find marginal density fns.

Soln $P((X,Y) \in A) = \iint_A f(x,y)\, dA$



$$= \int_{y=1/4}^{1/2} \int_{x=0}^{1/2} \frac{2}{5}(2x+3y)\, dx\, dy$$

$$= \int_{y=1/4}^{1/2} \frac{2}{5}\left(x^2+3xy\right)\Big|_{x=0}^{1/2} dy$$

$$= \frac{2}{5} \int_{y=1/4}^{1/2} \left(\frac{1}{4} + \frac{3}{2}y\right) dy$$

$$= \frac{2}{5}\left[\frac{y}{4} + \frac{3}{4}y^2\right]_{1/4}^{1/2}$$

$$= \frac{2}{5}\left[\frac{1}{8} + \frac{3}{16} - \frac{1}{16} - \frac{3}{64}\right]$$

$$= \frac{2}{5}\left[\frac{13}{64}\right] = \frac{13}{160}$$

ii) Marginal distributions

$$f_X(x) = \int_{y=-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{y=0}^{1} \frac{2}{5}(2x+3y)\,dy$$

$$= \frac{2}{5}\left(2xy + \frac{3}{2}y^2\right)\Bigg|_{y=0}^{1}$$

$$= \frac{2}{5}\left(2x + \frac{3}{2}\right)$$

$$= \frac{4x+3}{5}$$

$$f_Y(x) = \int_{x=-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_{x=0}^{1} \frac{2}{5}(2x+3y)\,dx$$

$$= \frac{2}{5}\left(x^2 + 3xy\right)\Bigg|_{x=0}^{1}$$

$$= \frac{2}{5}(1+3y)$$

Expected Value rule:

Let X and Y be continuous r.vs, let $g(X,Y)$ be a funtion of X and Y.

The expectation of $g(X,Y)$ is

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\,dx\,dy$$

Linearity of expection:

1) $E[aX+b] = aE[X]+b$

2) $E[X+Y] = E[X]+E[Y]$

Covariance and correlation

Let X and Y be continuous r.vs. Covariance of X and Y is

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

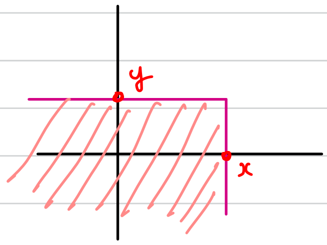
The Correlation between X and Y is,

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$$

Joint cumulative distribution fn.

Let X and Y be jointly continuous r.vs

$$F_{X,Y}(x,y) = P(X \le x,\ Y \le y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(x,y)\, dx\, dy$$



By fundamental thm,

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x\, \partial y}$$

Conditional probability density functions

Let X and Y be two continuous r.vs. with PDF $f_{X,Y}(x,y)$.

The conditional distribution of the r.v. Y given that $X=x$ is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad \text{if } f_X(x) > 0.$$

Since Conditional PDF $f_{Y|X}(y|x)$ is a PDF for all y in R_x the foll. properties are satisfied:

i) $f_{Y|X}(y|x) \ge 0$

ii) $\int_{y=-\infty}^{\infty} f_{Y|X}(y|x)\, dy = 1$

Multiplication rule:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_{Y|X}(y|x)$$
$$= f_Y(y) \cdot f_{X|Y}(x|y)$$

Conditional expectation and conditional variance

$$E[X \mid Y=y] = \int_{-\infty}^{\infty} x \, f_{X|Y}(x|y) \, dx$$

and $E[g(X) \mid Y=y] = \int_{-\infty}^{\infty} g(x) \, f_{X|Y}(x|y) \, dx$

$$\text{Var}(X \mid Y=y) = E[X^2 \mid Y=y] - \left(E[X \mid Y=y]\right)^2$$
$$= \int_x x^2 f_{X|Y}(x|y) \, dx - \left(\int_x x \, f_{X|Y}(x|y) \, dx\right)^2$$

Total probability and expectation thm

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dy$$
$$= \int_{-\infty}^{\infty} f_Y(y) \, f_{X|Y}(x|y) \, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx$$
$$= \int_{-\infty}^{\infty} f_X(x) \, f_{Y|X}(y|x) \, dx$$

$$E[X] = \int_{-\infty}^{\infty} f_Y(y) \, E[X \mid Y=y] \, dy$$

$$E[Y] = \int_{-\infty}^{\infty} f_X(x) \, E[Y \mid X=x] \, dx$$

pf: $E[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$

$$= \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f_Y(y) \, f_{X|Y}(x|y) \, dy\right) dx$$
$$= \int_{-\infty}^{\infty} f_Y(y) \left(\int_{-\infty}^{\infty} x \, f_{X|Y}(x|y) \, dx\right) dy$$
$$= \int_{-\infty}^{\infty} f_Y(y) \, E[X \mid Y=y] \, dy$$

Independence

For r.vs X and Y, the following properties are equivalent, and X and Y are independent.

1) $f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$ for all x and y

2) $f_{X|Y}(x|y) = f_X(x)$ for all x with $f_Y(y) \neq 0$

3) $f_{Y|X}(y|x) = f_Y(y)$ for all y with $f_X(x) \neq 0$

Independence and expectation

If r.vs X and Y are independent, then

$$E[XY] = E[X] \cdot E[Y]$$

$g(X)$ and $h(Y)$ are also independent:

$$E[g(X) \cdot h(Y)] = E[g(X)] \cdot E[h(Y)].$$

Note: In general, $E[XY] \neq E[X]\,E[Y]$

Independence and Variance

If X and Y are independent, then

$$Var(X+Y) = Var(X) + Var(Y)$$

Ex 2: Determine the value of c that makes the function

$$f(x,y) = c\,e^{-2x-3y}$$

a joint PDF over the range $0<x$ and $x<y$.

Determine the following:

a) $P(X<1, Y<2)$ b) $P(1<X<2)$ c) $P(Y>3)$

d) $P(X<2, Y<2)$ e) $E[X]$ f) $E[Y]$

g) Marginal Prob. distribution of X.

h) Conditional prob. distribution of Y given $X=1$

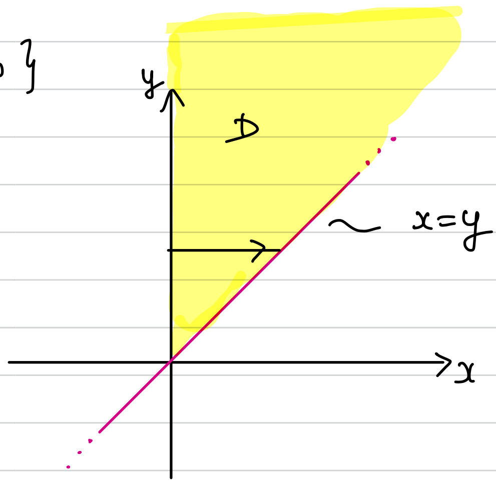
i) $E[Y \mid X=1]$

j) $P(Y<2 \mid X=1)$

k) Conditional prob. distribution of X given $Y=2$.

Soln: $f_{X,Y}(x,y) = c\, e^{-2x-3y}$ defind on D, where

$D = \{(x,y) \mid 0<x<y,\ 0 \le y < \infty\}$



To find c,

$$\iint_D c\, e^{-2x-3y}\, dA = 1$$

$$\Rightarrow \int_{y=0}^{\infty}\int_{x=0}^{y} c\, e^{-2x} \cdot e^{-3y}\, dx\, dy = 1$$

$$\Rightarrow \int_{y=0}^{\infty} c\, e^{-3y}\left(\frac{e^{-2y}}{-2} + \frac{1}{2}\right) dy = 1$$

$$\Rightarrow \left.\frac{c\, e^{-5y}}{10} + c\,\frac{e^{-3y}}{-6}\right]_0^{\infty} = 1$$

$$\Rightarrow -\frac{c}{10} + \frac{c}{6} = 1 \Rightarrow c = 15$$

Type 2

$D = \{(x,y) \mid 0 \le x \le y,\ 0 \le y \le \infty\}$

or

Type 1

$D = \{(x,y) \mid x \le y \le \infty,\ 0 \le x \le \infty\}$

a) $P(X<1, Y<2) = \iint_{D_1} f_{X,Y}(x,y)\, dA$

$$= \int_{x=0}^{1}\int_{y=x}^{2} 15\, e^{-2x-3y}\, dy\, dx$$

$$= \int_{x=0}^{1} 15\, e^{-2x} \cdot \left.\frac{e^{-3y}}{-3}\right]_{y=x}^{2} dx$$

$$= \int_{x=0}^{1} 15\, e^{-2x}\left(\frac{e^{-6}}{-3} + \frac{e^{-3x}}{3}\right) dx$$

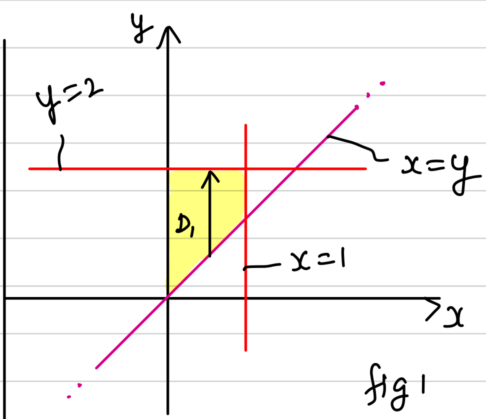


fig 1

Domain $= D_1$

$\{(x,y) \mid y \le x \le 2,\ x \le y \le 2\}$

$$= -5e^{-6}\left.\frac{e^{-2x}}{-2}\right]_{x=0}^{1} + 5\left.\frac{e^{-5x}}{-5}\right]_{x=0}^{1}$$

$$= \frac{5}{2}e^{-6}\left(e^{-2}-1\right) - \left(e^{-5}-1\right) = 0.9879$$

b) $P(1<X<2) = \int_{x=1}^{2}\int_{y=x}^{\infty} 15e^{-2x-3y}\,dy\,dx$

$$= \int_{x=1}^{2} 15e^{-2x}\left.\frac{e^{-3y}}{-3}\right]_{y=x}^{\infty} dx$$

$$= \int_{x=1}^{2} 5e^{-5x}\,dx$$

$$= \left.\frac{5e^{-5x}}{-5}\right]_{x=1}^{2}$$

$$= \left(e^{-10}-e^{-5}\right) = 0.0067$$

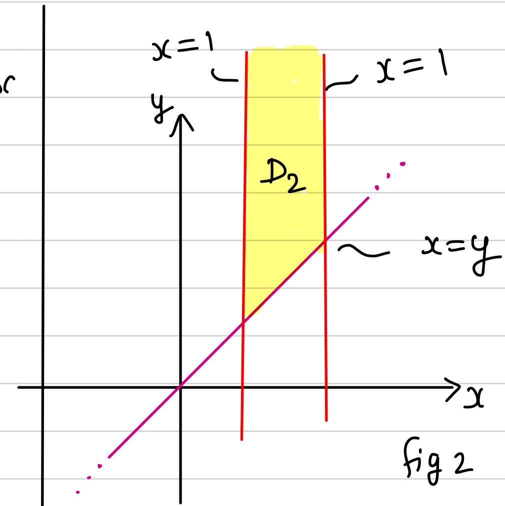


fig 2

$D_2 = \{(x,y) \mid x \le y \le \infty,\ 1 \le x \le 2\}$

c) $P(Y>3) = \int_{y=3}^{\infty}\int_{x=0}^{y} 15e^{-2x-3y}\,dx\,dy$

$$= \int_{y=3}^{\infty} 15e^{-3y}\left.\frac{e^{-2x}}{-2}\right]_{x=0}^{y} dy$$

$$= -\frac{15}{2}\int_{y=3}^{\infty}\left(e^{-5y}-e^{-3y}\right)dy$$

$$= -\frac{15}{2}\left[\frac{e^{-5y}}{-5} - \frac{e^{-3y}}{-3}\right]_{y=3}^{\infty}$$

$$= \frac{15}{2}\left[\frac{e^{-15}}{-5} + \frac{e^{-9}}{3}\right] = 0.000308$$

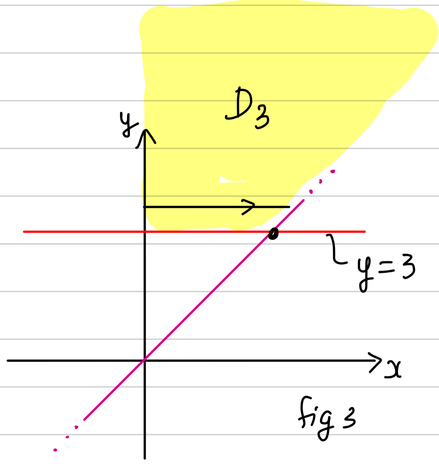


fig 3

$D_3 = \{(x,y) \mid 3 < y < \infty,\ 0 \le x \le y\}$

d) $P(X<2, Y<2) = \int_{x=0}^{2}\int_{y=x}^{2} 15e^{-2x-3y}\,dy\,dx$

$$= \int_{x=0}^{2} 5e^{-2x}\left(-e^{-6}+e^{-3x}\right)dx$$

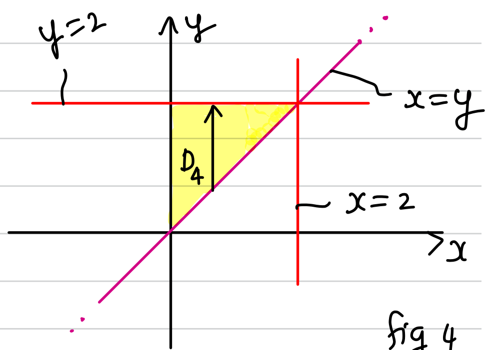


fig 4

$D_4 = \{(x,y) \mid x \le y \le 2,\ 0 \le x \le 2\}$

$$= 5\left[\frac{e^{-5x}}{-5} - e^{-6}\cdot\frac{e^{-2x}}{-2}\right]_0^2 \approx 0.9939$$

e) $E[X] = \iint_D x\, f_{X,Y}(x,y)\, dA$, where $D = \{(x,y) \mid x \le y \le \infty,\ 0 \le x \le \infty\}$

$$= 15\int_0^\infty\int_x^\infty x\, e^{-2x-3y}\, dy\, dx = 15\int_0^\infty x\, e^{-2x}\cdot\left.\frac{e^{-3y}}{-3}\right]_x^\infty dx$$

$$= 15\int_0^\infty x\, e^{-2x}\, \frac{e^{-3x}}{3}\, dx$$

$$= \frac{15}{3}\left[x\,\frac{e^{-5x}}{-5} - \frac{e^{-5x}}{25}\right]_0^\infty = \frac{1}{5}$$

f) $E[Y] = \iint_D y\, f_{X,Y}(x,y)\, dA$

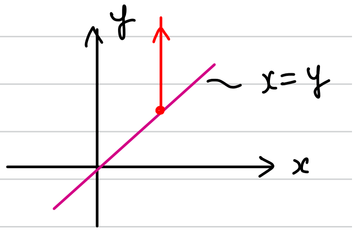
$$= 15\int_{x=0}^\infty\int_{y=x}^\infty y\, e^{-3y}\cdot e^{-2x}\, dy\, dx = 15\int_{x=0}^\infty e^{-2x}\left(y\,\frac{e^{-3y}}{-3} - \frac{e^{-3y}}{9}\right)_{y=x} dx$$

$$= 15\int_{x=0}^\infty e^{-2x}\left(\frac{x e^{-3x}}{3} + \frac{e^{-3x}}{9}\right) dx$$

$$= 5\left[x\,\frac{e^{-5x}}{-5} - \frac{e^{-5x}}{25}\right]_{x=0}^\infty + \frac{5}{3}\left[\frac{e^{-5x}}{-5}\right]_{x=0}^\infty$$

$$= \frac{1}{5} + \frac{1}{3} = \frac{8}{15}$$

g) Marginal Prob. distribution of X.



$x \le y \le \infty$

$$f_X(x) = \int_{y=-\infty}^\infty f_{X,Y}(x,y)\, dy$$

$$= \int_{y=x}^\infty 15\, e^{-2x-3y}\, dy = 5\, e^{-5x} \quad \text{for } x > 0$$

h) Conditional prob. distribution of Y given $X=1$

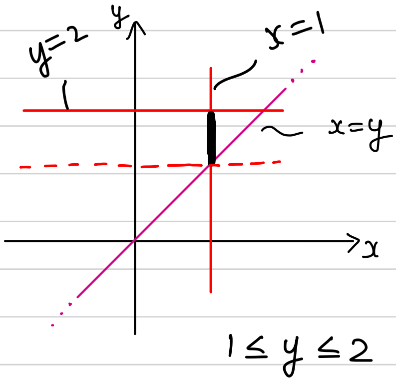
$$f_{Y|X}(y|1) = \frac{f_{X,Y}(1,y)}{f_X(1)} = \frac{15\, e^{-2-3y}}{5e^{-5}} = 3\, e^{-3y+3}, \quad y > 1$$

i) $E[Y|X=1] = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|1)\, dy = \int_{1}^{\infty} 3y\, e^{-3y+3}\, dy$

(At $x=1$, y varies from 1 to ∞)

$$= 3\left[\frac{y\, e^{-3y+3}}{-3} - \frac{e^{-3y+3}}{9}\right]_1^{\infty} = \frac{4}{3}$$

j) $P(Y<2|X=1) = \int_{y=1}^{2} f_{Y|X}(y|1)\, dy$



$$= \int_{y=1}^{2} 3\, e^{-3y+3}\, dy$$

$$= -e^{-3y+3}\Big]_1^2$$

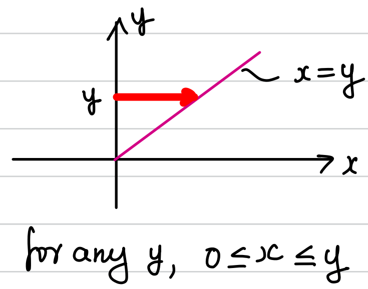
$$= -e^{-3} + 1 \simeq 0.9502$$

k) Conditional prob. distribution of X given $Y=2$.

$$f_Y(y) = \int_{x=-\infty}^{\infty} f_{X,Y}(x,y)\, dx = \int_{x=0}^{y} 15\, e^{-2x-3y}\, dx = -\frac{15}{2} e^{-5y} + \frac{15}{2} e^{-3y}$$

$$f_Y(2) = \frac{15}{2}\left(e^{-6} - e^{-10}\right)$$

$$f_{X|Y}(x|2) = \frac{f_{X,Y}(x,2)}{f_Y(2)} = \frac{\cancel{15}\, e^{-2x-6}}{\cancel{15}/2\,(e^{-6}-e^{-10})}$$



Ex3: Determine the value of c and the covariance and correlation for the joint PDF

$f_{X,Y}(x,y) = cxy$ over the range $0 < x < 2$ and $0 < y < x$

Soln: To find c

$$\int_{x=0}^{2}\int_{y=0}^{x} cxy\, dy\, dx = 1 \Rightarrow \int_{x=0}^{2} \frac{c}{2} x^3\, dx = 1 \Rightarrow c \cdot 2 = 1$$

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

$$E[X] = \int_{x=0}^{2}\int_{y=0}^{x} x\, f_{X,Y}(x,y)\, dy\, dx$$

$$= \int_{x=0}^{2}\int_{y=0}^{x} \frac{1}{2} x^2 y\, dy\, dx = \int_{x=0}^{2} \frac{1}{4} x^4\, dx = \frac{2^5}{20} = \frac{8}{5}$$

$$E[Y] = \int_{x=0}^{2}\int_{y=0}^{x} y\, f_{X,Y}(x,y)\, dy\, dx$$

$$= \int_{x=0}^{2}\int_{y=0}^{x} \frac{1}{2} x y^2\, dy\, dx$$

$$= \int_{x=0}^{2} \frac{1}{6} x^4\, dx = \frac{2^5}{30} = \frac{16}{15}$$

$$E[XY] = \int_{x=0}^{2}\int_{y=0}^{x} xy \cdot f_{X,Y}(x,y)\, dy\, dx$$

$$= \int_{x=0}^{2}\int_{y=0}^{x} \frac{1}{2} x^2 y^2\, dy\, dx = \frac{16}{9} \qquad \frac{4}{3} - \frac{256}{225} = \frac{44}{225}$$

Covariance, $\mathrm{COV}(X,Y) = \frac{16}{9} - \frac{8}{5} \times \frac{16}{15} = 0.071$

$E[X^2] = \frac{8}{3}$, $E[Y^2] = \frac{4}{3}$, $Var(X) = \frac{8}{75}$, $Var(Y) = \frac{44}{225}$

$$\therefore \rho(X,Y) = \frac{\mathrm{COV}(X,Y)}{\sqrt{Var(X)\, Var(Y)}} = 0.492$$

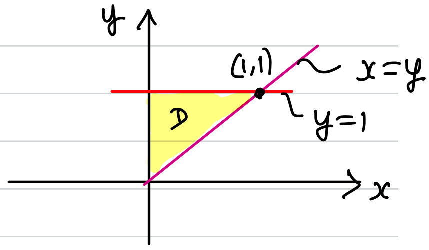
Ex4: The joint probability density function of two jointly continuous r.vs X and Y is

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

Find the marginal density fns of x and y. Also find the covariance between X and Y.

Soln: $f_{X,Y}(x,y) = 2$ in D, where $D = \{(x,y) \mid 0 < x < y,\ 0 < y < 1\}$

Marginal distribution of X,



$$f_X(x) = \int_y f_{X,Y}(x,y)\,dy$$

$$= \int_{y=x}^{1} 2\,dy, \qquad 0 < x < 1$$

$$\Rightarrow f_X(x) = 2(1-x), \qquad 0 < x < 1$$

Marginal distribution of Y,

$$f_Y(y) = \int_x f_{X,Y}(x,y)\,dx$$

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

$$\Rightarrow f_Y(y) = 2y, \qquad 0 < y < 1$$

$$E[X] = \iint_D x\, f_{X,Y}(x,y)\,dA \qquad \left(\text{or } \int_0^1 x\, f_X(x)\,dx \right)$$

$$= \int_{y=0}^{1}\int_{x=0}^{y} 2x\,dx\,dy$$

$$= \int_{y=0}^{1} y^2\,dy = 1/3$$

$$E[Y] = \iint_D y\, f_{X,Y}(x,y)\,dA \qquad \left(\text{or } \int_0^1 y\, f_Y(y)\,dy \right)$$

$$= \int_{y=0}^{1}\int_{x=0}^{y} 2y\,dx\,dy$$

$$= \int_{y=0}^{1} 2y^2\,dy = 2/3$$

$$E[XY] = \iint_D xy\, f_{X,Y}(x,y)\,dA = \int_{y=0}^{1}\int_{x=0}^{y} 2xy\,dx\,dy = \int_{y=0}^{1} y^3\,dy = 1/4$$

$\therefore \; Cov(X,Y) = E[XY] - E[X]\,E[Y] = \frac{1}{4} - \frac{1}{3}\cdot\frac{2}{3} = \frac{1}{36}.$

EX 5 : Verify that

$$f(x,y) = \begin{cases} e^{-x-y}, & x \geq 0 \text{ and } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is a density function of a joint probability distribution. Evaluate the following

i) $P(\frac{1}{2} < X < 2,\ 0 < Y < 4)$ ii) $P(X < 1)$, iii) $P(X \leq Y)$

iv) $P(X > Y)$ v) $P(X + Y \leq 1)$

Soln: $f(x,y) = e^{-x-y}$ defined on D, where

$$D = \{(x,y) \mid x \geq 0,\ y \geq 0\}$$

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

$\therefore$ It is a density function.

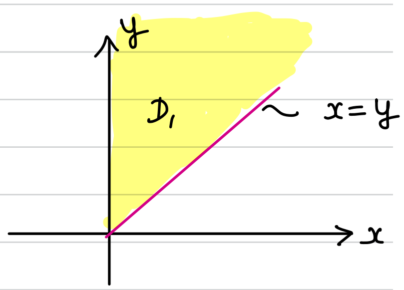
i) $P(\frac{1}{2} < X < 2,\ 0 < Y < 4) = \int_{x=\frac{1}{2}}^{2} \int_{y=0}^{4} e^{-x-y}\,dy\,dx$

$$= e^{-x}\Big]_{\frac{1}{2}}^{2}\; e^{-y}\Big]_0^4$$

$$= (e^{-2} - e^{-\frac{1}{2}})(e^{-4} - 1) = 0.4626$$

ii) $P(X < 1) = \int_{y=0}^{\infty} \int_{x=0}^{1} e^{-x-y}\,dx\,dy = (e^{-1} - 1)(-1) = 0.6321$

iii) $P(X \leq Y) = \iint_{D_1} e^{-x-y}\,dA$

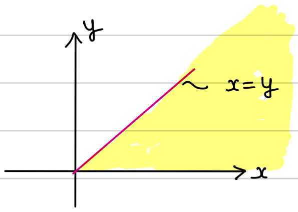


$$= \int_{x=0}^{\infty} \int_{y=x}^{\infty} e^{-x-y}\,dy\,dx$$

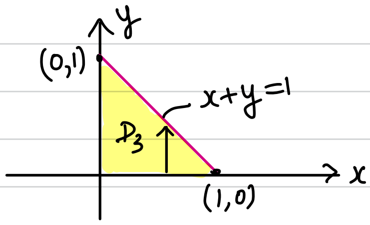
$$= \int_{x=0}^{\infty} e^{-x} \left[\frac{e^{-y}}{-1}\right]_x^\infty dx$$

$$= \int_{x=0}^{\infty} e^{-x} e^{-x} \, dx = \frac{e^{-2x}}{-2} \Big]_0^{\infty} = \frac{1}{2}$$

iv) $P(X > Y) = 1 - P(X \le Y) = \frac{1}{2}$



v) $P(X + Y \le 1) = \iint_{D_3} e^{-x-y} \, dA$



$$= \int_{x=0}^{1} \int_{y=0}^{1-x} e^{-x-y} \, dy \, dx$$

$$= \int_{x=0}^{1} e^{-x} \frac{e^{-y}}{-1} \Big]_{y=0}^{1-x} = \int_{x=0}^{1} e^{-x} \left(-e^{-1+x} + 1 \right) = \left[-e^{-1} x + \frac{e^{-x}}{-1} \right]_0^1$$

$$= -2e^{-1} + 1 = 0.2642$$

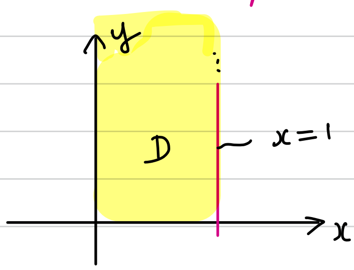
Ex6 : Find the constant K so that

$$f(x,y) = \begin{cases} K(x+1)e^{-y}, & 0 < x < 1, \ y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

is a joint prob. density fn. Are X and Y independent?

Soln: $f(x,y) = K(x+1)e^{-y}$ on D, where

$D = \{ (x,y) \mid 0 < x < 1, \ 0 < y < \infty \}$



Let $f(x,y)$ be PDF. Then

$$\iint_D f(x,y) \, dA = 1 \Rightarrow \int_{x=0}^{1} \int_{y=0}^{\infty} K(x+1)e^{-y} \, dy \, dx = 1$$

$$\Rightarrow \int_{x=0}^{1} K(x+1) \, dx = 1$$

$$\Rightarrow K \left(\frac{1}{2} + 1 \right) = 1 \Rightarrow K = \frac{2}{3}$$

Marginal distribution of X,

$$f_X(x) = \int_y f_{X,Y}(x,y)\,dy$$

$$= \int_0^\infty \frac{2}{3}(x+1)e^{-y}\,dy = \frac{2}{3}(x+1)$$

Marginal distribution of Y,

$$f_Y(y) = \int_x f_{X,Y}(x,y)\,dx$$

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

Clearly, $f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y)$

$\therefore$ X and Y are independent.

Ex7: The joint density for the r.v.s X and Y, where X is the unit temperature change and Y is the projection of spectrum shift that a certain atomic particle produces, is

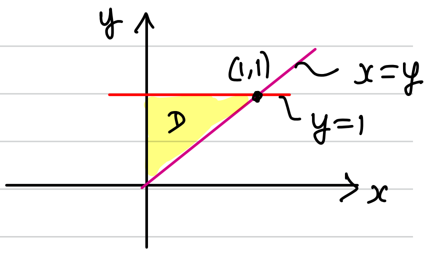
$$f(x,y) = \begin{cases} 10xy^2, & 0<x<y<1 \\ 0, & \text{elsewhere} \end{cases}$$

a) Find the marginal densities $f_X(x)$, $f_Y(y)$, and the conditional density $f_{Y|X}(y|x)$.

b) Find the probability that the spectrum shifts more than half of the total observations, given that the temperature is increased by 0.25 unit.

Soln: Given $f_{X,Y}(x,y) = 10xy^2$ defined on D, where

$D = \{(x,y) \mid 0<x<y,\ 0<y<1\}$



a) Marginal density fn of X,

$$f_X(x) = \int_y f_{X,Y}(x,y)\,dy = \int_{y=x}^{1} 10xy^2\,dy$$

$$\Rightarrow f_X(x) = \frac{10x}{3}(1-x^3), \quad 0<x<1$$

Marginal density fn of Y,

$$f_Y(y) = \int_x f_{X,Y}(x,y)\,dx = \int_0^y 10\,xy^2\,dx$$

$$\Rightarrow f_Y(y) = 5y^4, \quad 0<y<1$$

Conditional density,

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{10\,xy^2}{\frac{10}{3}x(1-x^3)} = \frac{3y^2}{1-x^3}, \quad 0<x<y<1$$

b) Therefore,

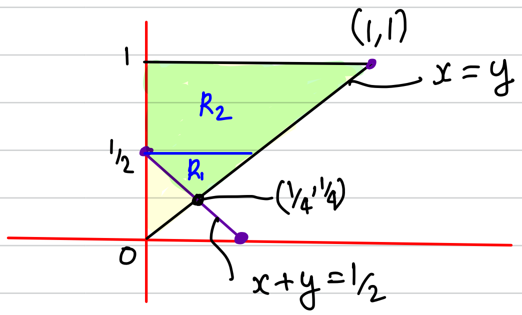
$$P(Y>1/2 \mid X=0.25) = \int_{y=1/2}^{1} f_{Y|X}(y|0.25)\,dy$$

$$= \int_{y=1/2}^{1} \frac{3y^2}{1-0.25^3}\,dy = \frac{64}{21}\left.\frac{y^3}{3}\right]_{1/2}^{1}$$

$$= \frac{64}{21}\left(\frac{1}{3}-\frac{1}{24}\right) = \frac{64}{21}\times\frac{7}{24} = \frac{8}{9}$$

Given:

$D = \{(x,y) \mid 0<x<y,\ 0<y<1\}$



$P(X+Y > 1/2)$

$D_3 = R_1 \cup R_2$

$R_1 = \left\{(x,y) \mid \frac{1}{4}<y<\frac{1}{2},\ \frac{1}{2}-y<x<y\right\}$

$R_2 = \{(x,y) \mid 1/2<y<1,\ 0<x<y$