Probability

Random experiments :

outcomes are unpredictable

eg. tossing a coin, rolling a die.

Trial :

a single performance of the random experiment

Sample space :

set of all outcomes of the random experiment

Rolling a die : $S = \{1, 2, 3, 4, 5, 6\}$

Tossing a coin : $S = \{H, T\}$

NOTE : a coin when tossed may land vertically. But we do not include that in the sample space because we focus only on the outcome that is of interest.

Rolling a die twice (or two dice) :

$$S = \{ (1,1), (1,2), (1,3), \ldots, (1,6),$$
$$(2,1), (2,2), (2,3), \ldots, (2,6),$$
$$\vdots$$
$$(6,1), (6,2), (6,3), \ldots, (6,6) \}$$

NOTE : The elements of the sample space are also called as sample points / realisation.

Event :

a collection of certain outcomes from the sample space.

i.e. an event is a subset of a sample space

for example, in the case of rolling a die twice the event of getting the sum is equal to '6' is

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

thus an event is any statement of conditions that defines the subset / a statement whose truth value is determined.

Sure event :

an event whose occurrence is inevitable whenever the experiment is performed.

for example it's a sure event to obtain a number between 1 and 6, on throwing a die.

ie event $A = \{1, 2, 3, 4, 5, 6\}$ is a sure event.

NOTE : sample space is a sure event in any random experiment.

Impossible event :

an event whose nonoccurrence is inevitable whenever the experiment is performed.

for any random experiment, null event (ϕ) is an impossible event.

Elementary event or atomic event :

event containing single outcome

Mutually exclusive events :

Two events A and B are mutually exclusive if

$A \cap B = \phi$

thus the two events cannot occur simultaneously.

$A = \{2, 4, 6\}$

$B = \{1, 3, 5\}$

$C = \{2, 3, 4, 5, 6\}$

A and B are mutually exclusive

A and C are not.

B and C are not.

a list of events $A_1, \dots A_n$ is mutually exclusive if $A_i \cap A_j = \phi$, $i \neq j$ ie no sample point is included in more than one event in the list.

$$\boxed{\bigcap_{i=1}^{n} A_i = \phi}$$

Collectively exhaustive events :

a list of events $A_1, A_2, \dots A_n$ is collectively exhaustive if $A_1 \cup A_2 \cup A_3 \dots \cup A_n = S$

$$\boxed{\bigcup_{i=1}^{n} A_i = S}$$

A collection of mutually exclusive and collectively exhaustive non empty events forms a partition of the sample space.

Probability :

measure of chance.
likelihood of an event to happen / occur.

3 approaches for finding probability.

i) Classical approach.
ii) Relative frequency approach.
iii) Axiomatic approach.

Classical Approach :

If an event E ~~has~~ happens in 'h' no. of favourable ways out of a total of 'n' possible ways all of which are equally likely, then the probability of E is given by

$$\boxed{P(E) = \frac{h}{n}}$$

it cannot handle events with an infinite number of possible outcomes.
it cannot handle events where each outcome is not equally-likely. (eg: unfair / biased coin, weighted die)

'Equally likely' is not ideal / realistic.
it is theoretical / an apriori approach. (without actually having to perform the experiment).

Relative frequency approach :

approach using past data / empirical approach.

An experiment is performed 'n' times, n is large, and event E happens 'h' no. of times, then the probability of E is estimated as $\boxed{P(E) = \lim_{n \to \infty} \frac{h}{n}}$

Example : Tossing coin :

n	h	P(H)
10	6	0.6
100	59	0.59
1000	613	0.613
10000	5897	0.5897

Limitations :

We can only attempt to make 'n' sufficiently large (as certain experiments can be costly / destructive).

may not attain a unique value. (approximate value).

Large is vague.

The main subject of probability theory is to develop tools to find probabilities for different events.

Axiomatic approach :

Let S be the sample space of a random experiment. To each event $A \subseteq S$, we assign a real no $P(A)$.

$P()$ is called a probability function if the following axioms are satisfied.

A_1 : $P(A) \geq 0$ for any $A \subseteq S$.

A_2 : $P(S) = 1$.

A_3 : For any list of m.e. events $A_1, A_2, \ldots$

$$P(\bigcup_i A_i) = P(A_1) + P(A_2) + \ldots$$

$$= \sum_i P(A_i)$$

The probability function $P : \text{Event space} \rightarrow \mathbb{R}$.

and $P(A)$ is the probability of the event A.

<u>Results</u> : (from the axioms)

<u>Result 1</u> : $\boxed{P(\phi) = 0.}$

<u>Proof</u> : $S \cup \phi = S$ (S, ϕ are m.e.)

$$P(S \cup \phi) = P(S)$$

by axiom 3 : $P(S) + P(\phi) = P(S)$

$$P(\phi) = 0$$

<u>Result 2</u> : $\boxed{P(\bar{A}) = 1 - P(A)}$

<u>Proof</u> : $A \cup \bar{A} = S$ ($A, \bar{A}$ are m.e)

$$P(A \cup \bar{A}) = P(S)$$

by axiom 3 : $P(A) + P(\bar{A}) = P(S)$

by axiom 2 : $P(A) + P(\bar{A}) = 1$

$$P(\bar{A}) = 1 - P(A).$$

Result 3 : If $\boxed{A \subseteq B}$, then $\boxed{\begin{array}{l} \text{i) } P(A) \le P(B) \\ \text{ii) } P(B-A) = P(B) - P(A). \end{array}}$

understanding $A \subseteq B$:

i) occurrence of A $\Rightarrow$ occurrence of B

ii) occurrence of B $\not\Rightarrow$ occurrence of A.

for example :

$A = \{3, 6\}$

$B = \{3, 4, 5, 6\}$

$A \subseteq B$.

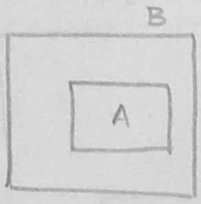
I outcome is 3.

$\therefore$ A has occurred ; B also has occurred

II outcome is 4.

$\therefore$ B has occurred ; A has not occurred

Proof :



$\Rightarrow$ $B = A \cup (B - A)$

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

by axiom 3 : $P(B) = P(A) + P(B-A)$

$P(B-A) = P(B) - P(A) \longrightarrow$ (ii)

Also $P(A) \le P(B)$ by axiom 1 ie. $P(B-A) \ge 0$.

Result 4 : $\boxed{0 \le P(A) \le 1}$.

Proof : $P(A) \ge 0$ from axiom 1 $\longrightarrow$ 1

W.R.t $A \subseteq S$

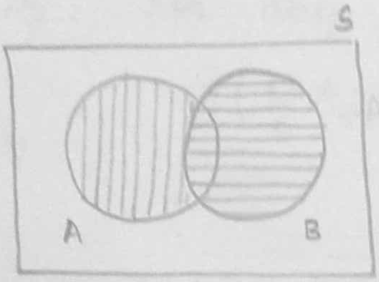
by result 3 : $P(A) \le P(S)$

by axiom 2 : $P(A) \le 1 \rightarrow$ 2

① & ② $\Rightarrow$ $0 \leq P(A) \leq 1$

Result 5 : Addition rule.

For any two events A, B $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$



Proof :
$$A \cup B = (A - B) \cup B$$
$$P(A \cup B) = P[(A - B) \cup B]$$

by axiom 3 : $P(A \cup B) = P(A - B) + P(B) \rightarrow ①$

$$A - B = A - (A \cap B).$$
$$P(A - B) = P[A - (A \cap B)], \quad (A \cap B) \subseteq A.$$
$$= P(A) - P(A \cap B), \text{ by result 3 (ii)}$$

$\therefore ① \Rightarrow P(A \cup B) = P(A) - P(A \cap B) + P(B)$

NOTE :

For three events A, B, C.

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

In general for n events $A_1, A_2, \ldots, A_n$.

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

Conditional probability :

$P(A|B)$ (read as probability of A given B).

Example.

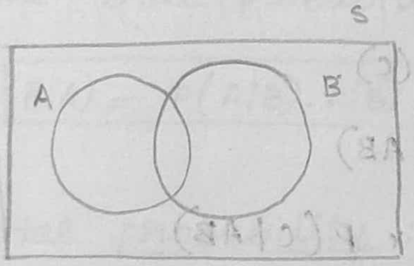
Tossing 2 coins

$S = \{HH, HT, TH, TT\}$

Suppose the o/c on the first coin is H, what is the probability that the second coin lands T?

$B = \{HH, HT\}$ becomes the reduced sample space

$P(A|B) = \frac{1}{2}$



$P(A) = \frac{\#(A)}{\#(S)}$

$^{*}P(A|B) = \frac{\#(A \cap B)}{\#(B)}$ provided $P(B) \neq 0$

$$P(A|B) = \frac{\frac{\#(A \cap B)}{\#(S)}}{\frac{\#(B)}{\#(S)}} \quad \text{from } *$$

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

Result 6: Let S be the sample space. A & B are events with $P(B) > 0$

i) $P(A|B) \geq 0$

ii) $P(S|B) = 1$

iii) For m.e events $A_1, A_2, \ldots$

$$P(\bigcup_i A_i \mid B) = \sum_i P(A_i | B)$$

ie. $P(A_1 \cup A_2 \cup \ldots \mid B) = P(A_1|B) + P(A_2|B) + \ldots$

iv) $P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2 | B)$

Multiplication rule :

Remember

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A \cap B) = P(A|B) \times P(B) \quad , \quad P(B) \neq 0$$

$$P(A \cap B) = P(B|A) \times P(A) \quad , \quad P(A) \neq 0 .$$

For 3 events A, B, C.

$$P(A \cap B \cap C) = P(ABC) = \cancel{P(AB|C) \times P(C)}.$$

$$= P(AB) \times P(C|AB)$$

$$= P(B) \times P(A|B) \times P(C|AB)$$

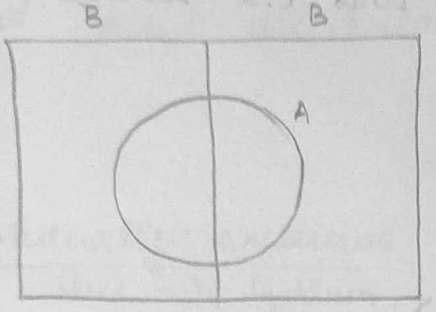
can be expanded in more than 1 way

$$\boxed{P(A_1 A_2 A_3) = P(A_2) \times P(A_1|A_2) \times P(A_3|A_1 A_2)}$$

In general :

$$P(A_1 A_2 A_3 \ldots A_n) = P(A_1) . P(A_2|A_1) . P(A_3|A_1 A_2) . P(A_4|A_1 A_2 A_3) \ldots P(A_n|A_1 A_2 \ldots A_{n-1}) .$$

Total probability and Bayes' rule



when finding $P(A)$ directly is not possible;

$A = (A \cap B) \cup (A \cap \bar{B})$ (mutually exclusive)

$P(A) = P(A \cap B) + P(A \cap \bar{B})$.

$= P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})$ (using multiplication rule)

Thus total probability of A:

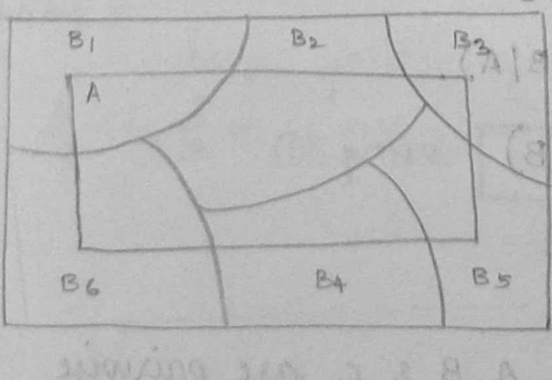
$$\boxed{P(A) = P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})}$$

Total probability:

Let $B_1, B_2, \ldots, B_n$ be a list of mutually exclusive and collectively exhaustive events in a sample space S with $P(B_i) \neq 0$. $\forall i$.

Then for any event A:

$$\boxed{P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \ldots + P(A|B_n) \cdot P(B_n)}$$



Bayes' rule :

$B_1, B_2, \ldots B_n$ is a list of m.e and c.e events.

A is any event.

Then $P(B_k/A) = \frac{P(A \cap B_k)}{P(A)}$.

multiplication rule

$$= \frac{P(B_k).P(A/B_k)}{P(B_1).P(A/B_1) + P(B_2).P(A/B_2) + \ldots + P(B_n).P(A/B_n)}$$

total rule

Thus

$$\boxed{P(B_k/A) = \frac{P(B_k).P(A/B_k)}{P(B_1).P(A/B_1) + P(B_2).P(A/B_2) + \ldots + P(B_n).P(A/B_n)}}$$

Independent events :

Two events A and B are independent if the occurrence or non occurrence of one event has no influence on the occurrence or non occurrence of the other event

ie $\boxed{P(A/B) = P(A)}$ and $\boxed{P(B/A) = P(B)} \rightarrow ①$

Result :

by multiplication rule

$P(A \cap B) = P(A).P(B/A)$.

$\boxed{P(A \cap B) = P(A).P(B)}$ using ①

NOTE :

$P(AB) = P(A).P(B)$

$P(BC) = P(B).P(C)$ $\Rightarrow$ A, B & C are pairwise independent.

$P(AC) = P(A).P(C)$

$P(ABC) = P(A).P(B).P(C) \Rightarrow A, B, C$ are independent.

and A, B, C are independent $\Rightarrow$ A, B and C are also pairwise independent

<u>Mutually exclusive vs. Independent events</u> :

A, B are m.e $\Rightarrow$ $P(A \cap B) = 0$.

A, B are ind. $\Rightarrow$ $P(A \cap B) = P(A).P(B)$

Independent events cannot be mutually exclusive untill both A and B have nonzero probabilities

<u>Problems</u> :

1. An elevator with 2 passangers stops at the 2^{nd}, 3^{rd}, 4^{th} floors. If it is equally likely that a passenger gets off at any of the 3 floors. What is the prob. they get off at different floors.

$a_i b_j$ - psg 1 gets off @ i^{th} floor
psg 2 gets off @ j^{th} floor
$i, j = 2, 3, 4$.

$S = \{a_2b_2, a_2b_3, a_2b_4, a_3b_2, a_3b_3, a_3b_4, a_4b_2, a_4b_3, a_4b_4\}$

$P(A) = \frac{6}{9} = \frac{2}{3}$

2. If 2 dice are thrown, what is the probability that the sum is

i) 8

ii) neither 7 nor 11.

i) $S = \{(1,1), \ldots, (1,6), (2,1), \ldots, (6,6)\}$

$n(S) = 36$.

$P(\text{sum is } 8) = \frac{5}{36}$

$P(\text{neither } 7 \text{ nor } 11) = 1 - P(\text{either } 7 \text{ or } 11)$

$= 1 - (P(\text{sum is } 7) + P(\text{sum is } 11))$ (m.e)

1,6
2,5 5,6
3,4 6,5
4,3
5,2
6,1

$= 1 - \frac{6}{36} - \frac{2}{36}$

$= \frac{36-8}{36}$

$= \frac{28}{36}$

$= \frac{7}{9}$

3. A card is drawn from a deck of well shuffled 52 cards. Find the probability that it is

i) either a spade or an ace.

Let A be the event of getting spade.
B be the event of getting an ace.

$n(S) = 52$. $n(A \cap B) = 1$

$n(A) = 13$

$n(B) = 4$.

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

$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52}$

$= \frac{16}{52}$

$= \frac{4}{13}$

4. The odds against A solving a problem are 4 to 3. The odds in favour of B solving the same problem are 7 to 5. What is the probability that the problem is solved if they both try independently.

A → A solving

B → B solving

$P(\bar{A}) = \frac{4}{7} \Rightarrow P(A) = \frac{3}{7}$

$P(B) = \frac{7}{12}$.

Given A, B are independent.

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

$= P(A) + P(B) - P(A).P(B)$

$= \frac{3}{7} + \frac{7}{12} - \frac{3}{\cancel{7}} \times \frac{\cancel{7}}{12}$

$= \frac{3}{7} + \frac{4}{12}$

$= \frac{36+28}{84}$

$= \frac{64}{84}$

2. If A and B are independent events th

i) $\bar{A}$ and $\bar{B}$ are also independent.

ii) $\bar{A}$ and B are also independent

iii) A and $\bar{B}$ are also independent.

i) <u>Proof</u>:

i) $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$

$= 1 - P(A \cup B)$

$= 1 - (P(A) + P(B) - P(A \cap B))$

$= 1 - P(A) - P(B) + P(A).P(B)$

$= 1 - P(A) - P(B)\{1 - P(A)\}$

$= [1 - P(A)][1 - P(B)]$

$= P(\bar{A}).P(\bar{B})$.

iii) $P(A \cap \bar{B}) = P(A - (A \cap B))$.

$= P(A) - P(A \cap B)$ (result 3)

$= P(A) - P(A).P(B)$.

$= P(A)[1 - P(B)]$

$= P(A).P(\bar{B})$.

ii) is similar to iii)

5. Three horses A, B, C are in a race. A is twice as likely to win as B. B is twice as likely to win as C. What are their respective chances of winning.

$P(A) = 2P(B)$
$P(B) = 2P(C)$.

A, B, C are mutually exclusive and collectively exhaustive. $\therefore P(A \cup B \cup C) = 1$.

$$P(A) + P(B) + P(C) = 1.$$

$$4P(C) + 2P(C) + P(C) = 1.$$

$$7 \cdot P(C) = 1$$

$$\boxed{P(C) = \frac{1}{7}}$$

$\therefore \boxed{P(B) = \frac{2}{7}}$.

$\boxed{P(A) = \frac{4}{7}}$.

6. Two boxes contain 3W, 4B and 4W, 3B balls. If a box is chosen and a ball is drawn from it. What is the prob. that it is a white ball.

$B_1, B_2 \rightarrow$ choosing Box 1, Box 2 respectively.

$P(B_1) = 1/2, \quad P(B_2) = 1/2$.

$A \rightarrow$ choosing white ball.

$P(A|B_1) = 3/7, \quad P(A|B_2) = 4/7$.

$\therefore P(A) = P(A/B_1).P(B_1) + P(A|B_2).P(B_2)$.

$$= \frac{3}{7}\cdot\frac{1}{2} + \frac{4}{7}\cdot\frac{1}{2}$$

$$= \frac{7}{14}$$

$$= 1/2$$

7. A box contains 5R and 3B. Another Box contains 4R and 5B balls. One of the boxes is chosen at random and two balls are drawn. What is the probability that one is red and other is blue?

$B_1, B_2 \rightarrow$ choosing Box$_1$, Box$_2$ respectively.

$A \rightarrow$ selecting 1 red, 1 blue balls.

$P(B_1) = 1/2$

$P(B_2) = 1/2$.

$$P(A|B_1) = \frac{5C_1 \cdot 3C_1}{8C_2}$$

$$= \frac{5 \times \cancel{3}^{3} \times \cancel{2}}{{}_4\cancel{8} \times \cancel{6} 7}$$

$$= \frac{15}{28}$$

$$P(A|B_2) = \frac{4C_1 \cdot 5C_1}{9C_2}$$

$$= \frac{4 \times 5 \times \cancel{2}}{8 \times 9}$$

$$= 20/36$$

$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2).$$

$$= \frac{15}{28} \cdot \frac{1}{2} + \frac{20}{36} \cdot \frac{1}{2}.$$

$$= \frac{15}{56} + \frac{20}{72}$$

$$= \frac{15 \times 9 + 20 \times 7}{504}$$

$$= \frac{135 + 140}{504} = \frac{275}{504}$$

8. A box contains 3 coins with a head on each side, four coins with a tail on each side and two fair coins. If one of these 9 coins are selected at random and tossed once, what is the probability that a head is obtained?

$B_1, B_2, B_3 \rightarrow$ selecting a twoheaded coin, a two-tailed coin, a fair coin respectively.

$P(B_1) = 3/9$

$P(B_2) = 4/9$

$P(B_3) = 2/9$.

$H \rightarrow$ tossing and obtaining a head.

$P(H|B_1) = 1$

$P(H|B_2) = 0$

$P(H|B_3) = 1/2$.

$\therefore P(H) = P(B_1).P(H|B_1) + P(B_2).P(H|B_2) + P(B_3).P(H|B_3)$

$= \frac{3}{9}.1 + \frac{4}{9}.0 + \frac{\cancel{2}}{9}.\frac{1}{\cancel{2}}$.

$= \frac{3}{9} + \frac{1}{9}$

$P(H) = \frac{4}{9}$.

9. A company exploys 3 ~~identical~~ analytical plants for the design and development of a particular products. Plans 1, 2, 3 are used 30%, 20% and 50% respectively The probabilities of a defective product of the plans are 0.01, 0.03 and 0.02.

i) If a random product was observed, what is the probability that its defective ?

ii) If it is found defective, which plan was most likely used and thus responsible ?

i)

$P(B_1) = \frac{30}{100} = 0.3$

$P(B_2) = 0.2$

$P(B_3) = 0.5$

$P(A|B_1) = 0.01$

$P(A|B2) = 0.03$

$P(A|B_3) = 0.02$

$\therefore P(A) = P(B_1).P(A|B_1) + P(B_2).P(A|B_2) + P(B_3).P(A|B_3)$

$= 0.3 \times 0.01 + 0.2 \times 0.03 + 0.5 \times 0.02$

$= 0.019$

ii)

$P(B_1|A) = \frac{P(B_1).P(A|B_1)}{P(A)}.$

$= \frac{0.3 \times 0.01}{0.019}$

$= 0.157$

$P(B_2/A) = 0.315$

$P(B_3|A) = 0.526.$

Plan 3 is responsible.

10. There are 10 urns, each of three contains 1 W and 9 B Balls, each of another three contains 9 W and 1 B balls and each of the remaining four contains 5 W and 5 B balls. One of the urns is selected at random and a ball chosen from it is found to be white. What is the probability that it was urn containing 1 W and 9 B balls?

$B_1, B_2, B_3 \rightarrow$ urn type I, II, III.

$P(B_1) = 3/10$, $P(B_2) = 3/10$, $P(B_3) = 4/10$.

$P(W/B_1) = 1/10$, $P(W/B_2) = 9/10$, $P(W/B_3) = 5/10$

$$P(B_1/W) = \frac{P(B_1) \cdot P(W/B_1)}{P(B_1) \cdot P(W/B_1) + P(B_2) \cdot P(W/B_2) + P(B_3) \cdot P(W/B_3)}$$

$$= \frac{3/100}{3/100 + 27/100 + 20/100}$$

$$= \frac{3/100}{50/100}$$

$$= \frac{3}{50}$$

: Probability - Basic concepts tutorial.

12.

$P(T_1) = 0.40$

$P(T_0) = 0.60$

$P(R_0|T_0) = 0.90$, $P(R_1|T_0) = 0.10$

$P(R_1|T_1) = 0.95$, $P(R_0|T_1) = 0.05$

i)

$P(R_1) = P(T_1).P(R_1|T_1) + P(T_0).P(R_1|T_0)$.

$= 0.40 \times 0.95 + 0.60 \times 0.10$.

$= 0.44$

ii) $P(T_1|R_1) = \dfrac{P(T_1).P(R_1|T_1)}{P(R_1)}$.

$= \dfrac{0.40 \times 0.95}{0.44}$

$= 0.8636$

iii) $P(\text{error}) = P(T_0 \cap R_1) + P(T_1 \cap R_0)$.

$= P(T_0).P(R_1|T_0) + P(T_1).P(R_0|T_1)$

$= 0.60 \times 0.10 + 0.40 \times 0.05$

$= 0.08$.

1. $n(S) = 4+8$. (reduced sample space)

$P(G) = \dfrac{n(G)}{n(S)}$

$= \dfrac{4}{12}$.

$= \dfrac{1}{3}$.

can also be solved using conditional probability.

2.

$S = \{HH, HT, TH, TT\}$

A : Head appears in the first toss $P(A) = \frac{2}{4} = 1/2$

B : " " " " second toss $P(B) = \frac{2}{4} = 1/2$

C : same outcome on both toss $P(C) = \frac{2}{4} = 1/2$.

$P(A \cap B \cap C) = 1/4$.

$P(A).P(B).P(C) = 1/8$.

$P(A \cap B \cap C) \neq P(A).P(B).P(C)$.

$\therefore$ A, B and C are not independent events.

NOTE :

A, B and C are pairwise independent (verify!)

3. $S = \{HH, HT, TH, TT\}$

A : atmost one head $P(A) = 3/4$

B : one head and one tail $P(B) = 2/4$

$P(A \cap B) = 2/4$.

$P(A)P(B) = 6/16 = 3/8$.

$P(A \cap B) \neq P(A).P(B)$.

$\therefore$ A, B are not independent events.

4. $P(A) = 1/3$, $P(B) = 3/4$, $P(A \cup B) = 11/12$.

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

$= 1/3 + 3/4 - 11/12$

$= \frac{4+9-11}{12} = \frac{2}{12} = \frac{1}{6}$

2 $\therefore P(A|B) = \frac{P(A\cap B)}{P(B)}$

$$= \frac{1}{\cancel{6}_3} \times \frac{\cancel{4}^2}{3}$$

$$= \frac{2}{9}$$

5. $P(A) = \frac{75}{100} = \frac{3}{4}$ $\qquad P(B) = \frac{80}{100} = \frac{4}{5}$

$P(\bar{A}) = 1/4$ $\qquad P(\bar{B}) = 1/5$

contradiction between Alice and Bob can occur in 2 ways.

Case 1 : A speaks truth but B lies.

$$P(C_1) = P(A \cap \bar{B})$$

$$= P(A).P(\bar{B}).$$

$$= \frac{3}{4} \cdot \frac{1}{5}$$

$$= \frac{3}{20}.$$

Case 2 : A lies but B speaks the truth

$$P(C_2) = P(\bar{A} \cap B)$$

$$= P(\bar{A}).P(B)$$

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

$$= \frac{4}{20}.$$

$P(C) = P(C_1) + P(C_2)$ (m.e)

$$= \frac{3}{20} + \frac{4}{20}$$

$$= \frac{7}{20}$$

6. $P(A|B) = 0.2$.

$P(A|\bar{B}) = 0.3$

$P(B) = 0.8 \quad \therefore P(\bar{B}) = 0.2$

$P(A) = P(B).P(A|B) + P(\bar{B}).P(A|\bar{B})$

$= 0.8 \times 0.2 + 0.2 \times 0.3$

$= 0.22$

7. A and B are mutually exclusive and exhaustive events.

$A \cup B = S$. ($\because$ collectively exhaustive)

$P(A \cup B) = P(S)$.

$P(A \cup B) = 1$

$P(A) + P(B) = 1$ (mutually exclusive).

$P(A) + 2P(A) = 1$

$3P(A) = 1$

$P(A) = 1/3$

8. $P(A) = 0.6$ $\quad P(A \cap B) = 0.2$

$P(B) = 0.7$.

W.k.t. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

$= 0.6 + 0.7 - 0.2$.

$= 1.3 - 0.2$

$= 1.1$. (not possible).

2 10.

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

A : obtaining no heads.

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

$\therefore$ The argument is invalid.

9.

$$P(\text{sum } 6) = \frac{n(\{(1,5), (2,4), (3,3), (4,2), (5,1)\})}{36}$$

$$P(S_6) = 5/36 \quad \therefore P(\bar{S_6}) = 31/36$$

$$P(\text{sum } 7) = \frac{n(\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\})}{36}$$

$$P(S_7) = 6/36 = 1/6 \quad \therefore P(\bar{S_7}) = 5/6$$

given : A begins the game.

$$P(A \text{ wins}) = P(S_6) + P(\bar{S_6}).P(\bar{S_7}).P(S_6)$$
$$+ P(\bar{S_6}).P(\bar{S_7}).P(\bar{S_6}).P(\bar{S_7}).P(S_6)$$
$$+ \dots$$
$$= P(S_6)\left[1 + P(\bar{S_6}).P(\bar{S_7}) + [P(\bar{S_6}).P(\bar{S_7})]^2 + [P(\bar{S_6}).P(\bar{S_7})]^3 + \dots\right]$$

$$= \frac{P(S_6)}{1 - P(\bar{S_6}).P(\bar{S_7})}.$$

$$= \frac{5/36}{1 - 31/36 \times 5/6} = \frac{5/36}{1 - \frac{155}{216}}$$

$= \frac{5}{36} \times \frac{216}{61}$ 6

$= \frac{30}{61}$

11. A_k = The number k appears on the die A.

$P(A_1) = P(B_1) = 0.1$

$P(A_2) = P(B_2) = 0.1$

$P(A_3) = P(B_3) = 0.3$

$P(A_4) = P(B_4) = 0.3$

$P(A_5) = P(B_5) = 0.1$

$P(A_6) = P(B_6) = 0.1$.

$P(\text{sum is } 7) = P((A_1 \cap B_6) \cup (A_2 \cap B_5) \cup (A_3 \cap B_4) \cup (A_4 \cap B_3) \cup (A_5 \cap B_2) \cup (A_6 \cap B_1))$

$= P(A_1).P(B_6) + P(A_2).P(B_5) + P(A_3).P(B_4) + P(A_4).P(B_3) + P(A_5).P(B_2) + P(A_6).P(B_1)$.

$= 2(0.1 \times 0.1 + 0.1 \times 0.1 + 0.3 \times 0.3)$

$= 2(0.01 + 0.01 + 0.09)$

$= 2(0.11)$

$= 0.22$

2

Random variables

A random variable is a function that assigns to each outcome $w \in S$ a real number. (or) to each event $A \subseteq S$ a real number.

A real valued function $X : \Omega \to R$ is a random variable if for each interval $I \subseteq R$, $\{w, X(w) \in I\}$ is an event.

Eg: 3 satellites are launched into space. Let the random variable X be the number of satellites that go into the orbit.

$S = \{FFF, FFS, FSF, SFF, SSF, SFS, FSS, SSS\}$

$X(FFF) = 0$ $\quad$ $X(SSF) = 2$
$X(FFS) = 1$ $\quad$ $X(SFS) = 2$
$X(FSF) = 1$ $\quad$ $X(FSS) = 2$
$X(SFF) = 1$ $\quad$ $X(SSS) = 3$

A person tosses coin until head appears.

$S = \{H, TH, TTH, TTTH, \ldots\}$

Let the random variable X be the number of times the coin is tossed to get head.

$X(H) = 1$
$X(TH) = 2$.
$\vdots$

$\therefore$ space $(X) = \{1, 2, 3, 4, \ldots\}$

A person is waiting for a call from his/her friend starting from 6:00 am 3/8/21.

S = {All time instances from 6:00 a.m of 3/8/21}

Let the random variable X be the amount of time he/she waits for the call.

Space (X) = [0, ∞)

A person is throwing a dart on a circular dart board of radius 1 m.

S = {All the points inside the circle}

Let the random variable X be the distance by which the person misses the bull's eye (centre).

space (X) = [0, 1)

Types of random variable :

i) Discrete	ii) Continuous
# has discrete values	# has continuous values
# takes countable number of values.	# takes any value in the interval of its range
# Eg 1, 2, 3.	# Eg between 3 and 4.

Another example : Tossing 2 dice

X : sum of the 2 dice
Y : max of the 2 dice.

X_i : o/c of the i^{th} die

$\therefore \quad X = X_1 + X_2$

2 Distribution function (df) or cumulative density function (cdf)

For a random variable X, the function $F : \mathbb{R} \rightarrow [0,1]$ defined by $F(x) = P[X \leq x]$ for $x \in \mathbb{R}$ is called the df / cdf.

Eg: $S = \{HHH, HHT, HTH, THH, THT, TTH, HTT, TTT\}$

X : number of heads.

$P(X = 0) = 1/8$

$P(X = 1) = 3/8$

$P(X = 2) = 3/8$

$P(X = 3) = 1/8$

$\therefore F(1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8}$

$F(1.2) = \frac{4}{8}$

$F(2.7) = \frac{7}{8}$

$F(3) = 1$

$F(5) = 1$

$F(-3) = 0$

Properties

i) $F(x)$ is real valued non-decreasing function ($x < y$, then $F(x) \leq F(y)$)

ii) $F(x)$ is right continuous function ($\lim_{x \to a^+} F(x) = F(a)$

iii) $\lim_{x \to -\infty} F(x) = 0$

iv) $\lim_{x \to \infty} F(x) = 1$

When dealing with a random variable X, for any a, b where $a < b$, we want to find the probabilities for one or more of the following events.

$[X \le a]$ $\quad [a < x < b]$

$[X < a]$ $\quad [a < x \le b]$

$[X > a]$ $\quad [a \le x < b]$

$[X \ge a]$ $\quad [a \le x \le b]$

$[x = a]$

$[x \ne a]$

Events	Probability.
i) $[x \le a]$	$F(a)$
ii) $[X < a]$	$F(a^-)$
iii) $[X > a]$	$1 - F(a)$
iv) $[X \ge a]$	$1 - F(a^-)$
v) $[X = a]$	$F(a) - F(a^-)$
vi) $[X \ne a]$	$1 - F(a) + F(a^-)$
vii) $[a < x < b]$	$F(b^-) - F(a)$
viii) $[a < x \le b]$	$F(b) - F(a)$
ix) $[a \le x < b]$	$F(b^-) - F(a^-)$
x) $[a \le x \le b]$	$F(b) - F(a^-)$

2 Probability distribution

+ probability mass function (pmf) : X is discrete
probability density function (pdf) : X is continuous

PMF :

f is the pmf of a discrete random variable X, if $f(x_i) = P(X = x_i) \ \forall x_i \in$ Space (X).

NOTE : i) $f(x_i) \geq 0$

ii) $\sum_{x_i} f(x_i) = 1$

If pmf 'f' is known, F can be found

$$F(x) = P[X \leq x] = \sum_{x_i < x} f(x_i)$$

PDF :

For a continuous random variable X, we talk about the probability that X takes a value in an interval. The function f is called the pdf of a continuous random variable X if.

$$\int_a^b f(x)\,dx = P[a < X < b]$$

If pdf 'f' is known, F can be found

$$F(a) = P[X \leq a] = \int_{-\infty}^{a} f(x)\,dx$$

NOTE :

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

ii) F(x) is a probability measure while f(x) is not a probability measure unless it's multiplied by an infinitesimal Δx to yield $f(x)\Delta x$

$$f(x) = \lim_{\Delta x \to 0^+} \frac{P[x < X < x + \Delta x]}{\Delta x}$$

$$= \lim_{\Delta x \to 0^+} \frac{F(x + \Delta x) - F(x)}{\Delta x}$$

$$\boxed{f(x) = F'(x)}$$

1. The pdf of a random variable x is given by

$$f(x) = \begin{cases} k(2x - x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

i) Find 'k'.

ii) Find $P[X > 1]$.

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

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

$$0 + k\left[x^2 - \frac{x^3}{3}\right]_0^2 = 1$$

$$k\left(4 - \frac{8}{3} - 0\right) = 1$$

$$k\,\frac{4}{3} = 1$$

$$\boxed{k = \frac{3}{4}}$$

ii)
$$P[X > 1] = \frac{3}{4}\int_{1}^{2} (2x - x^2)\,dx.$$

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

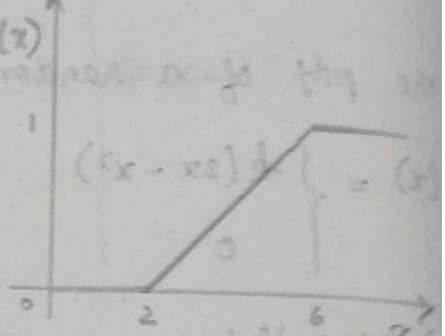
$$= \frac{3}{4}\left[\frac{4}{3} - 1 + \frac{1}{3}\right]$$

$$= \frac{3}{4} \times \frac{2}{8}$$

$$\boxed{P(X>1) = \frac{1}{2}}$$

2. The cdf of a r.v. X is

$$F(x) = \begin{cases} 0, & x < 2 \\ k(x-2), & 2 \le x < 6 \\ 1, & x \ge 6 \end{cases}$$



i) find k.

ii) find $P[X>4]$

iii) Find $P[3 \le X \le 5]$.

i)

$$F(6^-) = F(6).$$

$$k(6-2) = 1$$

$$k(4) = 1$$

$$\boxed{k = 1/4}$$

ii) $P[X>4] = 1 - P[X \le 4]$

$$= 1 - F(4)$$

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

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

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

iii) $P[3 \le X \le 5] = F(5) - F(3^-)$

$$= \frac{1}{4}(3) - \frac{1}{4}(1)$$

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

3. Two dice are rolled. Let X denote the absolute value of the difference of the o/c s. Obtain the pmf of X.

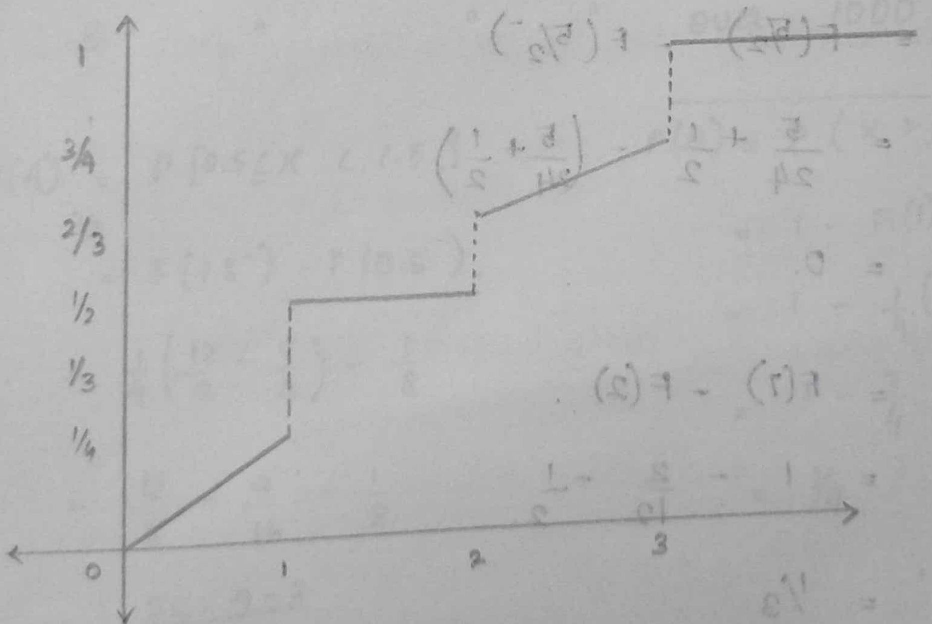
space (X) = $\{0, 1, 2, 3, 4, 5\}$

$X = x$	0	1	2	3	4	5
$P(X = x)$	$1/6$	$10/36$	$8/36$	$6/36$	$4/36$	$2/36$

4. The cdf of a r.v is given by

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ x/4 & , \quad 0 \leq x < 1 \\ 1/2 & , \quad 1 \leq x < 2 \\ \frac{x}{12} + \frac{1}{2} & , \quad 2 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

i) sketch the graph of F.



ii)

$P[X<2] = F(2^-)$

$= 1/2$

$P[X=2] = F(2) - F(2^-)$

$= \frac{2}{12} + \cancel{\frac{1}{2}} - \cancel{\frac{1}{2}}$

$= 1/6$.

$P[1<X<3] = F(3^-) - F(1^-)$.

$= \frac{\cancel{3}^{1}}{\cancel{12}_{4}} + \frac{1}{2} - \frac{1}{4}$.

$= 1/2$

$P[X>3/2] = 1 - P[X \le 3/2]$

$= 1 - F(3/2)$.

$= 1 - 1/2$

$= 1/2$

$P[X=5/2] = F(5/2) - F(5/2^-)$

$= \frac{5}{24} + \frac{1}{2} - \left(\frac{5}{24} + \frac{1}{2}\right)$

$= 0$.

$P[2<X \le 7] = F(7) - F(2)$.

$= 1 - \frac{2}{12} - \frac{1}{2}$.

$= 1/3$

4. The sales of a convenience store on a randomly selected day is a r.v with cdf (in 1000's of USD)

$$F(x) = \begin{cases} 0 & x < 0 \\ x^2/2 & 0 \le x < 1 \\ k(4x - x^2) & 1 \le x < 2 \\ 1 & x \ge 2. \end{cases}$$

i) Suppose the store's total sales on any given day is less than 2000. Find the value of k.

$$P[X < 2] = 1.$$

$$F(2^-) = 1$$

$$k(8-4) = 1$$

$$4k = 1$$

$$k = 1/4$$

ii) Let A : Tomorrow's sales is b/w 500 and 1500.

B : " " " over 1000

$$P(A) = P[0.5 \le X < 1.5]$$

$$= F(1.5^-) - F(0.5^-).$$

$$= \frac{1}{4}\left(\frac{12}{2} - \frac{9}{4}\right) - \frac{1}{8}$$

$$= \frac{12}{8} - \frac{9}{16} - \frac{1}{8}$$

$$= \frac{24-9-2}{16}$$

$$= \frac{13}{16}$$

$$P(B) = P(X > 1)$$

$$= 1 - F(1)$$

$$= 1 - \frac{1}{4}(4-1)$$

$$= 1 - \frac{3}{4}$$

$$= 1/4$$

$$P(A \cap B) = P[1 < X < 3/2]$$

$$= F(3/2^-) - F(1)$$

$$= \frac{1}{4}\left(\frac{15}{4}\right) - \frac{3}{4}$$

$$= \frac{15}{16} - \frac{3}{4}$$

$$= \frac{15 - 12}{16}$$

$$= \frac{3}{16}$$

$$P(A) \cdot P(B) = \frac{13}{16} \cdot \frac{1}{4}$$

$$= \frac{13}{64}$$

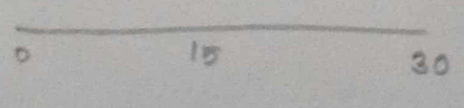
$$P(A) \cdot P(B) \neq P(A \cap B)$$

5. Suppose that a bus arrives at a bus stop every day between 10:00 and 10:30 AM. Let X denote the waiting time. Find the distribution function (cdf) of F.

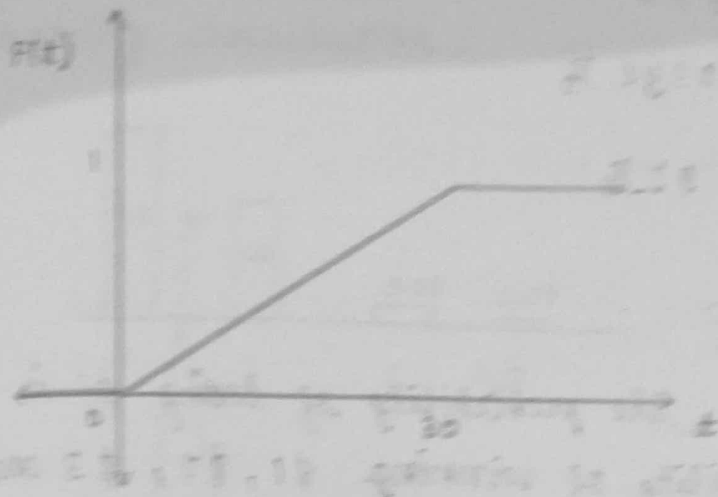
space $(X) = [0, 30)$

To find : $F(t)$, the distribution function

$$F(t) = \begin{cases} 0 & t < 0 \\ \frac{t}{30} & 0 \leq t < 30 \\ 1 & t \geq 30 \end{cases}$$



all time instances in $[0, 30)$ are equally likely $\therefore F(t) = \frac{t}{30}$



5. Let X be a point selected at random from $(0,1)$. Find the distribution function (cdf) of $Y = \frac{X}{1+X}$

<u>cdf of X:</u>

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$

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

$$= P\left[\frac{X}{1+X} \le y\right]$$

$$= P[X \le y + Xy]$$

$$= P[X - Xy \le y]$$

$$= P\left[X \le \frac{y}{1-y}\right]$$

$$= F_X\left(\frac{y}{1-y}\right)$$

$$\therefore \quad F_Y(y) = \frac{y}{1-y} \qquad 0 < \frac{y}{1-y} < 1$$

$$\text{or } \quad 0 < y < 1-y$$

$$0 < 2y < 1$$

$$0 < y < 1/2$$

$$\therefore F_Y(y) = \begin{cases} 0 & y \le 0 \\ \frac{y}{1-y} & 0 < y < 1/2 \\ 1 & y \ge 1/2 \end{cases}$$

Mean and Variance :

In a casino game, the probability of losing \$1 is 0.6 and the probability of winning \$1, \$2, \$3 are 0.3, 0.08, 0.02 respectively.

If this game is played n times, the gain is

$$(-1)(0.6)n + 1(0.3)n + 2(0.08)n + 3(0.02)n$$

$$= -0.08n.$$

In average there is a loss of \$ 0.08 per game.

Let X denote the r.v. gain space $(X) = \{-1, 1, 2, 3\}$

$$E(X) = -0.08. = \mu$$

$E(X)$ - expected value of X.

NOTE : expected value need not be a value that the random variable assumes.

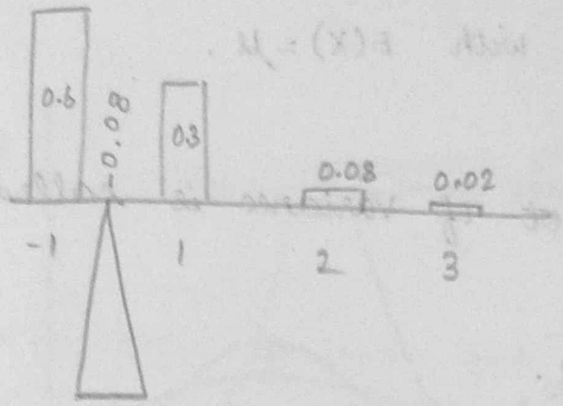
Thus for a random variable X, the expected value of X is defined by.

$$E(X) = \begin{cases} \sum x\, p(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x f(x)\, dx & \text{if } X \text{ is continous} \end{cases} \quad *$$

$E(X)$ is the weighted average of all values of X in which the weight assigned to each $x \in X$ is $p(X)$.

* provided the sum or integral exists.

This mean E(X) is a measure of the centre of the probability distribution.



centre of gravity of unit mass

In a physical sense, it is the centre of mass.

Properties :

i) $E(K) = K$ where K is constant (constant random variable)

ii) $E(Kx) = K\,E(x)$

iii) $E(X+Y) = E(X) + E(Y)$

iv) $E(X-Y) = E(X) - E(Y)$.

Variance :

Consider the random variables.

$X = 0$ with prob 1

$$Y = \begin{cases} -1 & \text{with prob } 1/2 \\ 1 & \text{" " } 1/2 \end{cases}$$

$$Z = \begin{cases} -10 & \text{" " } 1/4 \\ 0 & \text{" " } 1/2 \\ 10 & \text{" " } 1/4 \end{cases}$$

$E(X) = E(Y) = E(Z) = 0$

∴ Expected value by itself is not capable of providing adequate information.

∴ Expected value only gives the centre of distribution.

Thus another measure called variance is introduced.

For a random variable x with $E(X) = \mu$.

$$var(X) = E(X-\mu)^2.$$

NOTE : i) The square root of variance is called standard deviation.

$$SD(X) = \sigma_X = \sqrt{var(X)}.$$

ii) The standard deviation has the same unit as the random variable.

Alternate formula :

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

$$\begin{aligned} var(X) &= E[(X-\mu)^2] \\ &= E[X^2 - 2X\mu + \mu^2]. \\ &= E(X^2) - 2\mu E(X) + \mu^2. \\ &= E(X^2) - 2[E(X)]^2 + [E(X)]^2. \\ &= E(X^2) - [E(X)]^2. \end{aligned}$$

Properties :

i) $var(K) = 0$ where K is constant.

ii) $var(KX) = K^2 \, var(X)$.

iii) $var(X+Y) = Var(X) + Var(Y) + 2\underbrace{[E(XY) - E(X)E(Y)]}_{\text{covariance }(X,Y)}$

NOTE :

If X and Y are independent,

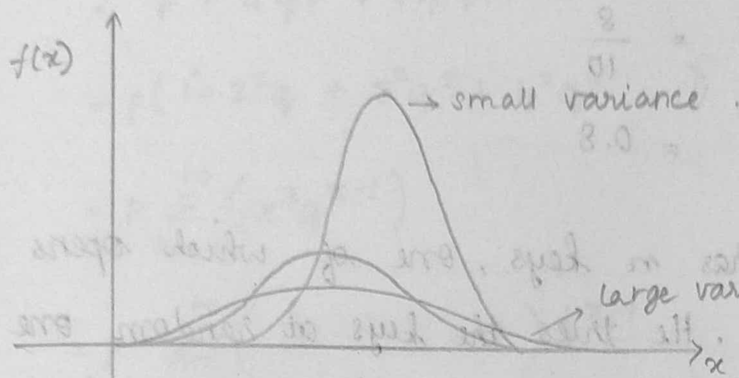
$$cov(X,Y) = 0.$$

$\therefore var(X+Y) = var(X) + var(Y)$ if X and Y are independent.

In a physical sense, it is the moment of inertia.

NOTE :

i) larger variance → more spread and pdf is wide and flat.



ii) small variance → less spread and pdf is narrow and tall.

Problems :

1. A box contains five balls, 2 of which are marked \$1 another 2 are marked \$5 and the remaining 1 is marked \$15. select 2 balls at random and win the sum of the amounts. To play the game, you pay \$10. What is the expected gain ?

(1) (1) (5) (5) (15)

(1,1), (1,5), (1,15), (5,5), (5,15).

Let the random variable X denote the gain.

	1,1	1,5	1,15	5,5	5,15
$X=x$	-8	-4	6	0	10
$P(x)$	$\frac{2C_2}{5C_2} = \frac{1}{10}$	$\frac{2C_1 \cdot 2C_1}{5C_2} = \frac{4}{10}$	$\frac{2C_1}{5C_2} = \frac{2}{10}$	$\frac{2C_2}{5C_2} = \frac{1}{10}$	$\frac{2C_1}{5C_2} = \frac{2}{10}$

Expected gain = $E(X) = \sum x P(x)$

$$= \frac{-8}{10} - \frac{16}{10} + \frac{12}{10} + 0 + \frac{20}{10}$$

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

$$= 0.8$$

2. A drunken man has n keys, one of which opens the door of his house. He tries the keys at random one by one and independently. compute the mean and variance of the no. of trials to open the door if wrong keys are

i) not eliminated.

ii) eliminated.

The random variable X is the number of trials to open the door.

Case 1:

$X = x$	1	2	3	...	∞
$P(x)$	p	qp	q^2p	...	

where $p = 1/n$ and $q = 1 - 1/n$.

$1 + x + x^2 + \dots = \frac{1}{1-x}$

$1 + 2x + 3x^2 + \dots = \frac{1}{(1-x)^2}$

$\frac{1}{2}(1.2 + 2.3x + \dots) = \frac{1}{(1-x)^3}$

$E(X) = \sum x P(x)$

$$= p + 2qp + 3q^2p + \dots$$

$$= p(1 + 2q + 3q^2 + \dots)$$

$$= p\left(\frac{1}{(1-q)^2}\right)$$

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

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

$E(X) = n.$

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

$E(X^2) = \sum x^2 P(x)$.

$= p + 4qp + 9q^2p + \dots$

$= p(1^2 + 2^2q + 3^2q^2 + 4^2q^3 \dots)$.

$= p\sum_{x=1}^{\infty}(x^2q^{x-1})$.

$= p\sum_{x=1}^{\infty}((x-1)x + x)q^{x-1}$

$= p\{\sum(x-1)xq^{x-1} + \sum xq^{x-1}\}$

$= p\{(0 + 1.2q + 2.3q^2 + \dots) + (1 + 2q + 3q^2 + \dots)\}$

$= p\{q(1.2 + 2.3q + 3.4q^2 + \dots) + \frac{1}{(1-q)^2}\}$

$= p\{q\frac{2}{(1-q)^3} + \frac{1}{(1-q)^2}\}$.

$= p\{\frac{2q}{p^3} + \frac{1}{p^2}\}$.

$= \frac{2q}{p^2} + \frac{1}{p}$.

$var(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$

$= \frac{2q + p - 1}{p^2}$

$= \frac{2q - q}{p^2}$

$= \frac{q}{p^2}$.

$= \frac{1 - 1/n}{1/n^2} = \frac{n-1}{\cancel{n}} \times n^2 = n(n-1)$.

Case 2 :

$X=x$	1	2	3	. . .	n
$P(x)$	$1/n$	$\frac{n-1}{n}\cdot\frac{1}{n-1}$	$\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\frac{1}{n-2}$		
	$1/n$	$1/n$	$1/n$	. . .	$1/n$
	p	p	p	. . . .	p

$$E(X) = \Sigma x P(x)$$

$$= p + 2p + 3p + \dots np$$

$$= p(1+2+3+\dots n)$$

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

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

$$E(X^2) = \Sigma x^2 P(x)$$

$$= p + 4p + 9p + \dots n^2 p$$

$$= p(1+2^2+3^2+\dots+n^2)$$

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

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

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

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{2(n+1)(2n+1) - 3(n+1)^2}{12}$$

$$= \frac{(n+1)[2(2n+1) - 3(n+1)]}{12}$$

$$= \frac{(n+1)(4n+2-3n-3)}{12}$$

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

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

Results :

For a continuous random variable x with cdf F and pdf f

$$\boxed{E(X) = \int_0^\infty [1-F(t)]\,dt - \int_0^\infty F(-t)\,dt.}$$

or

$$\boxed{E(X) = \int_0^\infty P(X > t)\,dt - \int_0^\infty P(X \le -t)\,dt.}$$

NOTE :

This result is true even if x is not continuous.

When x is a non-negative random variable.

$$E(X) = \int_0^\infty [1-F(t)]\,dt \quad \text{and} \quad E(X^2) = 2\int t[1-F(t)]\,dt$$

or

$$E(X) = \int_0^\infty P(X > t)\,dt.$$

Problem :

1. The distribution function cdf of a r.v X is given by $F(X) = \begin{cases} 1 - 16/x^2 & x \ge 4 \\ 0 & x < 4 \end{cases}$

Find E(X). Show that var(X) does not exist.

$$E(X) = \int_0^\infty [1 - F(t)]\,dt.$$

$$= \int_4^\infty [\cancel{1} - \cancel{1} + \frac{16}{t^2}]\,dt + \int_0^4 (1-0)\,dt$$

$$= 16\left[-\frac{1}{t}\right]_4^{\infty} + [t]_0^4$$

$$= 16\left(0-(-1/4)\right) + (4-0)$$

$$= \frac{\cancel{16}^{4}}{\cancel{4}} + 4$$

$$E(X) = 8$$

$$E(X^2) = 2\int_0^{\infty} t\,[1-F(t)]\,dt$$

$$= 2\int_4^{\infty} t\,\frac{16}{t^2}\,dt$$

$$= 2\int_4^{\infty} \frac{16}{t}\,dt$$

$$= 32\,[\log t]_4^{\infty}$$

= does not exist as $\log\infty$ does not exist

∴ Variance does not exist.

2. A box contains 10 discs of radius 1, 2, ... 10 resp. Find the expected value of the circumference of a randomly selected disc.

X be the circumference.

$X = x$	2π	4π	...	20π
$P(x)$	$1/10$	$1/10$	...	$1/10$

$$E(X) = \sum x\,P(x)$$

$$= 2\pi/10 + 4\pi/10 + \dots 20\pi/10$$

$$= \frac{2\pi}{10}(1+2+3+\dots+10)$$

$$= \frac{2\pi}{\cancel{10}} \times \frac{\cancel{10}\times 11}{2}$$

$$E(X) = 11\pi$$

3. The time it takes for a student to finish an exam has pdf

$$f(x) = \begin{cases} 6(x-1)(2-x) & 0 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find the mean and SD.

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

$$= \int_0^2 x\, 6(x-1)(2-x)\,dx .$$

$$= \int_0^2 6x(2x - x^2 - 2 + x)\,dx$$

$$= \int_0^2 (18x^2 - 6x^3 - 12x)\,dx .$$

$$= \overset{6}{\cancel{18}}\left[\frac{x^3}{\cancel{3}}\right]_0^2 - \overset{3}{\cancel{6}}\left[\frac{x^4}{\underset{2}{\cancel{4}}}\right]_0^2 - \overset{6}{\cancel{12}}\left[\frac{x^2}{\cancel{2}}\right]_0^2$$

$$= 6(8-0) - 3(8-0) - 6(4-0).$$

$$= 48 - 24 - 24$$

$$= 0.$$

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

$$= \int_0^2 x^2\, 6(x-1)(2-x) .$$

$$= 6\int_0^2 (-x^4 + 3x^3 - 2x^2)\,dx$$

$$= 6\left[\frac{-x^5}{5} + \frac{3x^4}{4} - \frac{2x^3}{3}\right]_0^2$$

$$= 6\left[-\frac{32}{5} + 12 - \frac{16}{3} - (0)\right]$$

$$= 6\left(\frac{-96+180-80}{15}\right)$$

$$= \frac{24}{15}$$

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

$$= \frac{24}{15}$$

$$SD = \sqrt{Var(X)}$$

$$= \sqrt{\frac{24}{15}}$$

$$= \frac{2\sqrt{6}}{\sqrt{15}}$$

Special Probability Distributions:

1. Discrete uniform
2. Bernoulli
3. Binomial
4. Poisson
5. Geometric

Discrete uniform distribution

A random variable x that takes values $a, a+1, \ldots, b$ is discrete uniform random variable if it assigns equal probabilities for all these values.

parameter

pmf : $$f(x) = \begin{cases} \frac{1}{b-a+1} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

cdf : $$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a+1}{b-a+1} & a \le x \le b \\ 1 & x > b \end{cases}$$

mean : $E(X) = \frac{a+b}{2}$

variance : $Var(X) = \frac{n^2-1}{12}$ where $n = b-a+1$

Bernoulli distribution

A Bernoulli experiment is an experiment that has only 2 outcomes (success and failure).

If a Bernoulli experiment is repeated it is called Bernoulli trials.

Let $P(\text{success}) = p$ parameter
$P(\text{failure}) = q = 1-p$.

The random variable X that takes values

$$X = \begin{cases} 1 & \text{if the outcome is success} \\ 0 & \text{if the outcome is failure} \end{cases}$$

is a Bernoulli random variable.

pmf : $f(x) = \begin{cases} p & x=1 \\ q & x=0 \end{cases}$

mean : $E(X) = p$

Variance : $Var(X) = p - p^2 = p(1-p) = pq$

Binomial Distribution

consider a sequence of n 'iid' Bernoulli trials with $P(\text{success}) = p$ in each trial.

iid - independent and identical.

The random variable X is number of successes in n trials.

Space $(X) = \{0, 1, 2, 3, \ldots, n\}$

pmf : $f(x) = {}^nC_x\, p^x q^{n-x}$, $x = 0, 1, 2, \ldots n$

parameters : n and p

mean : $E(X) = np$

variance : $var(X) = npq$

Geometric Distribution

Consider a sequence of 'iid' Bernoulli trials, with the number of trials not fixed and $P(\text{success}) = p$ in each trial.

Here the random variable X is the number of trials for success.

pmf : $f(x) = q^{x-1} p$, $x = 1, 2, 3, \ldots$

parameter : p

Mean : $E(X) = 1/p$

variance : $var(X) = q/p^2$

Memoryless property of geometric distribution

$$P[X > m+n \mid X > m] = P[X > n]$$

This property is also called the forgetfulness property. It means that a given probability distribution is independent of its history.

$$P[X > n] = \sum_{x=n+1}^{\infty} p(x)$$

$$= \sum q^{x-1} p \ , \ x = n+1, n+2, \ldots$$

$$= q^n p + q^{n+1} p + \ldots$$

$= p(q^n + q^{n+1} + \dots)$

$= pq^n (1 + q + q^2 + \dots)$

$= \cancel{p} q^n \frac{1}{\cancel{1-q}}$

$= q^n$.

$\therefore P[X > r] = q^r$

$$P[X > m+n \mid X > m] = \frac{P[X > m+n,\ X > m]}{P[X > m]}$$

$$= \frac{P[X > m+n]}{P[X > m]}$$

$$= \frac{q^{m+n}}{q^m}$$

$$= q^n$$

Poisson Distribution :

Consider a random variable X that denotes the number of occurrences of an event in a specified interval of time / space.

Examples

\# number of messages received in 1 minute.

\# number of misprints in a page

pmf : $f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, 2, \dots$

parameter : $\lambda > 0$ is the mean

NOTE :

\# Poisson Distribution is a limiting form of Binomial distribution where $n \to \infty$ and p is close to 0 or 1.

* PD is used to model the probability distribution for rare events occurring in a specified boundary.

mean :

$$E(X) = \Sigma x\, P(x)$$

$$= \Sigma x \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \Sigma \frac{\lambda^x}{(x-1)!}$$

$$= \lambda e^{-\lambda} \Sigma \frac{\lambda^{x-1}}{(x-1)!}, \quad x = 1, 2, \ldots$$

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

$$= \lambda e^{-\lambda} e^{\lambda}.$$

$$= \lambda.$$

variance :

$$E(x^2) = \Sigma x^2 p(x)$$

$$= \Sigma x^2 \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \Sigma (x^2 - x + x) \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \left\{ \Sigma x(x-1) \frac{\lambda^x}{x!} + \Sigma x \frac{\lambda^x}{x!} \right\}$$

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

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

$$= e^{-\lambda} \left\{ \lambda^2 e^{\lambda} + \lambda e^{\lambda} \right\}$$

$$= \lambda^2 + \lambda$$

$$var(X) = E(X^2) - [E(X)]^2$$
$$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2}$$
$$= \lambda.$$

NOTE :

In Poisson Distribution $E(X) = Var(X)$.

Problems :

1. If 4 dice are rolled what is the probability that at most one 6 appears.?

X : no. of times 6 appears.

$p = 1/6$

$q = 5/6.$

B.D with $n = 4$

$$P(x) = 4C_x p^x q^{4-x}, \quad n = 0, 1, 2, 3, 4.$$

$$P[X \le 1] = P[X = 0] + P[X = 1]$$
$$= 4C_0 (1/6)^0 (5/6)^4 + 4C_1 (1/6)^1 (5/6)^3.$$
$$= \left(\frac{5}{6}\right)^4 + \frac{4}{6}\left(\frac{5}{6}\right)^3.$$
$$= \left(\frac{5}{6}\right)^3 \left(\frac{5}{6} + \frac{4}{6}\right)$$
$$= \frac{125}{\underset{24}{\cancel{216}}} \times \frac{\cancel{9}}{6}$$
$$= \frac{125}{144}$$

2. The probability of an item produced by a machine is defective is 0.05. What is the probability that the machine will produce a defective item in its 6^{th} run?

Let x be the number of runs to see a defective item.

$X \sim G.D(p = 0.05)$

$P(X=6) = (1-p)^5 p$

$= 0.95^5 \times 0.05.$

$= 0.03869$

3. The probability that an applicant for driving license will pass the road test on any given trial 0.8. Find the probability that he will finally pass test

i) on the 4^{th} trial.

ii) in fewer than 4 trials.

Let x be the number of trials to pass the test.

$X \sim G.D(0.8)$.

i) $P[X=4] = 0.2^3 . 0.8$

$= 0.0064$

ii) $P[X<4] = 1 - P[X>3]$

$= 1 - (0.2)^3$

$= 0.992$

4. A person has 2 taxis. The no. of demands for taxi on each day follows Poisson Distribution with mean 1.5. Find the probability that

i) neither taxi is used

ii) some demands are refused.

Let x be number of demands for taxi.

$X \sim PD(1.5)$

$$f(x) = \frac{e^{-1.5}(1.5)^x}{x!}$$

i) $P[X=0] = \frac{e^{-1.5} \times 1}{1}$

$= e^{-1.5}$

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

$= 1 - e^{-1.5} - 1.5e^{-1.5} - \frac{1.5^2 e^{-1.5}}{2}$

$= 1 - e^{-1.5}(1 + 1.5 + 1.125)$

$= 1 - e^{-1.5}(3.625)$

5. In a certain factory manufacturing razor blades, there is a small chance of 1/500 for any blade to be defective. The blades are supplied in packets of 10. Use poisson distribution to find the number of packets containing

i) no defective

ii) one defective

iii) two defective

blades in a consignment of 10000 packets.

Let x be the number of defective blades in a pack of 10.

$X \sim PD(\lambda = 0.02)$ $\lambda = \text{mean} = np = 10 \times \frac{1}{500} = 0.02$

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

i) $P[X=0] = e^{-0.02} = 0.9802$

ii) $P[X=1] = e^{-0.02}(0.02) = 0.0196.$

iii) $P[X=2] = \frac{e^{-0.02}(0.02)^2}{2!} = 0.000196.$

In a consignment of 10000 packets, no. of packets containing

i) no defective blade : 9802

ii) one defective blade : 196

iii) two defective blades : 1.96 ≃ 2.

6. Guests arrived in a hotel, in accordance with a poisson process at a rate of 5 per hour. Suppose that for the last 10 minutes no guest has arrived. What is the probability that the next one will arrive in less than 2 minutes ?

Let $X(t)$ be the number of times an event has occurred by time 't'.

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

$\lambda = 5$ per hour.

$\therefore$ no. of arrivals $\sim$ P.D(5)

$$P[X(2/60) = n] = \frac{e^{-5\times 2/60}\,(5\times 2/60)^n}{n!}$$

$$= \frac{e^{-1/6}\,(1/6)^n}{n!}, \quad n = 0, 1, 2, \ldots$$

P[next guest in < 2min] = 1 − P[next guest in more then 2 min]

$$= 1 - P[X(t) = 0]$$

$$= 1 - e^{-1/6}$$

Special continuous distribution

1. continuous uniform
2. Exponential
3. Weibull
4. Normal

Continuous uniform distribution.

If X is a point selected at random from the interval (a, b), the X is called a uniform random variable over (a, b).

pdf : $f(x) = \begin{cases} \dfrac{1}{b-a} & x \in (a, b) \\ 0 & \text{otherwise} \end{cases}$

cdf : $F(x) = \begin{cases} \dfrac{x-a}{b-a} & x \in (a, b) \\ 0 & x \in (-\infty, a] \\ 1 & x \in [b, \infty) \end{cases}$

$$P[k_1 < X < k_2] = \frac{k_2 - k_1}{b-a}$$

mean :

$$E(X) = \int x f(x)\,dx$$

$$= \int_a^b \frac{x}{b-a}\,dx.$$

$$= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

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

$$= \frac{(b+a)\cancel{(b-a)}}{2\cancel{(b-a)}}$$

$$= \frac{a+b}{2}$$

variance :

$$E(x^2) = \int x^2 f(x)\,dx$$

$$= \int_a^b \frac{x^2}{b-a}\,dx$$

$$= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$= \frac{1}{3(b-a)}(b^3-a^3)$$

$$= \frac{\cancel{(b-a)}(b^2+ab+a^2)}{3\cancel{(b-a)}}$$

$$= \frac{b^2+ab+a^2}{3}$$

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

$$= \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4}$$

$$= \frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12}$$

$$= \frac{b^2 - 2ab + a^2}{12}$$

$$= \frac{(b-a)^2}{12}.$$

Problems :

1. It takes a student a random time between 20 and 27 minutes to walk from hostel to college. If he has class at 8.30 AM and he leaves hostel at 8.07 AM. Find the probability that he reaches his class on time.

Let x be the time taken to reach college (a random point between 20 & 27).

$X \sim$ uniform $(20, 27)$

$$P[\text{reaches on time}] = P[X \leq 23]$$

$$= F(23)$$

$$= \frac{23-20}{27-20}$$

$$= \frac{3}{7}$$

2. Buses arrive at a bus stop every 15 minutes starting at 7:00 AM ie 7:00, 7:15, 7:30 and so on. If a passenger at time that is uniformly distributed b/w 7 and 7:30 AM. Find the probability that they wait

i) less than 5 minutes

ii) more than 10 minutes.

What is the average time spent waiting ?

Let x be the arrival time of the passenger past 7:00 a.m

$\therefore$ $X \sim$ uniform $(0, 30)$

i) P[passenger waits < 5 min]

$$= P[10 < X \leq 15] + P[25 < X \leq 30]$$

$$= \frac{15-10}{30-0} + \frac{30-25}{30-0}$$

$$= \frac{5}{30} + \frac{5}{30}$$

$$= \frac{10}{30}$$

$$= \frac{1}{3}$$

ii) P[waits more than 10 min]

$$= P[0 < X < 5] + P[15 < X < 20]$$

$$= \frac{5-0}{30-0} + \frac{20-15}{30-0}$$

$$= \frac{5}{30} + \frac{5}{30}$$

$$= \frac{10}{30}$$

$$= \frac{1}{3}$$

Let Y be the time spent waiting

$\therefore$ $Y \sim$ uniform $(0, 15)$

$$\therefore \quad E(Y) = \frac{15+0}{2}$$

$$= 7.5 \text{ minutes.}$$

Exponential Distribution

time b/w emergency arrivals in a hospital

time b/w catastrophic events

time b/w failures of components aka lifetime.

Let X be the time b/w consecutive occurrences of an event.

pdf : $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.

parameter : $\lambda > 0$

time & frequency relationship

exp. r.v :	time b/w consecutive events
poisson r.v :	no. of arrivals per unit time

cdf : $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$

mean :

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

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

$$\boxed{\int_0^{\infty} e^{-ax} x^n dx = \frac{n!}{a^{n+1}}}$$

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

$$= \lambda \frac{1!}{\lambda^2}$$

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

variance :

$$E(X^2) = \int x^2 f(x) dx$$

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

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

$$= \lambda \frac{2!}{\lambda^3} = \frac{2}{\lambda^2}$$

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

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

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

<u>For exponential distribution:</u>

$$\text{variance} = (\text{mean})^2.$$

<u>Memoryless property of exponential distribution:</u>

$$P[X > t+s \mid X > s] = P[X > t]$$

$$P[X>t] = \int_t^\infty f(x)\,dx$$

$$= \int_t^\infty \lambda e^{-\lambda x}\,dx.$$

$$= \cancel{\lambda}\left[\frac{e^{-\lambda x}}{-\cancel{\lambda}}\right]_t^\infty$$

$$= [-0 - (-e^{-\lambda t})]$$

$$\boxed{P[X>t] = e^{-\lambda t}.}$$

$$P[X > t+s \mid X > s] = \frac{P[X > t+s,\ X > s]}{P[X>s]}$$

$$= \frac{P[X > t+s]}{P[X > s]}$$

$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}}.$$

$$= e^{-\lambda(t+s)+\lambda s}.$$

$$= e^{-\lambda t}$$

Problems :

1. Suppose the time one spends in a bank follows ED. with a mean of 10 minutes. What is the probability that a customer will spend more than 15 minutes in the bank?

What is the probability that she will spend more than 15 minutes given she is still in the bank after 10 minutes?

Let x be the amount of time spent in the bank (in minutes).

$\therefore$ $X \sim ED(\lambda = 1/10)$.

$$P[X > 15] = e^{-0.1(15)}$$

$$= e^{-1.5}$$

$$P[X > 5+10 \mid X > 10] = P[X > 5]$$

$$= e^{-0.1(5)}$$

$$= e^{-0.5}$$

6. Guests arrived in a hotel .... refer (poisson distribution)

Let x be the time taken for the arrival of next guest.

$$\lambda = 5/hr = \frac{5}{60} / min.$$

$\therefore$ $X \sim ED(\lambda = 1/12)$

$$P[X > 10+2 \mid X > 10] = P[X > 2]$$

$$= e^{-1/12(2)}$$

$$= e^{-1/6}$$

$$\therefore P(\text{required}) = 1 - e^{-1/6}$$

Weibull distribution :

x : time to failure of a component.

more like reliability.

$$R(t) = P[X > t]$$

If the pdf of x is

$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}} \quad ; \; x \geq 0$$

then X follows Weibull distribution.

parameters : α, β

α : scale parameter

β : shape parameter

NOTE : when $\beta = 1$, Weibull distribution becomes exponential distribution.

mean :

$$E(X^r) = \int_0^{\infty} x^r f(x)\,dx$$

$$= \int_0^{\infty} x^r \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}\,dx.$$

Take $\boxed{\begin{array}{l} u = \alpha x^{\beta} \\ du = \alpha\beta x^{\beta-1}\,dx \end{array}}$ $\boxed{\begin{array}{l} x^{\beta} = u/\alpha \\ x = (u/\alpha)^{1/\beta} \end{array}}$

$$= \int_0^{\infty} \left(\frac{u}{\alpha}\right)^{r/\beta} e^{-u}\,du.$$

$$= \frac{1}{(\alpha)^{r/\beta}} \int_0^{\infty} u^{r/\beta} e^{-u}\,du$$

$$E(X^r) = \frac{1}{\alpha^{r/\beta}} \Gamma\left(\frac{r}{\beta}+1\right)$$

$$\therefore \; E(X) = \frac{1}{a^{1/\beta}} \Gamma\left(\frac{1}{\beta}+1\right)$$

variance :

$$E(X^2) = \frac{1}{\alpha^{2/\beta}} \Gamma\left(\frac{2}{\beta}+1\right).$$

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

$$= \frac{1}{\alpha^{2/\beta}} \Gamma\left(\frac{2}{\beta}+1\right) - \frac{1}{\alpha^{2/\beta}}\left[\Gamma\left(\frac{1}{\beta}+1\right)\right]^2$$

$$P[X>t] = \int_t^\infty f(x)\,dx$$

$$= \int_t^\infty \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} dx$$

$$\boxed{\begin{array}{l} u = \alpha x^\beta \\ du = \alpha\beta x^{\beta-1}\,dx. \\ x = t. \Rightarrow u = \alpha t^\beta \\ x = \infty \Rightarrow u = \infty. \end{array}}$$

$$= \int_{\alpha t^\beta}^\infty e^{-u}\,du$$

$$= \left[\frac{e^{-u}}{-1}\right]_{\alpha t^\beta}^\infty$$

$$= -(0 - e^{-\alpha t^\beta})$$

$$\boxed{P[X>t] = e^{-\alpha t^\beta}} \qquad \boxed{F(t) = 1 - e^{-\alpha t^\beta}}$$

Failure rate / Hazard rate $z(t)$

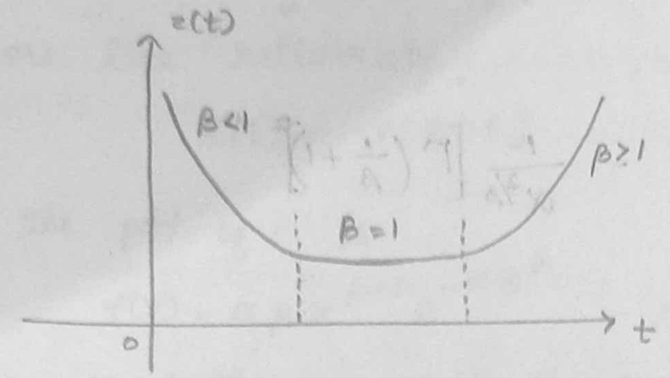
$$z(t) = \alpha\beta t^{\beta-1}$$

i) $\beta < 1$ $\Rightarrow$ $z(t)$ is decreasing
the components strengthen over time

ii) $\beta = 1$ $\Rightarrow$ $z(t) = \alpha$, constant
becomes exponential distribution in which memoryless property prevails.

iii) $\beta > 1$ $\Rightarrow$ $z(t)$ is increasing

the components wear out over time.



β is called shape parameter because it gives the shape of the distribution.

Problem :

1. Suppose the lifetime (in hrs) of a battery follows Weibull distribution with $\alpha = 0.1$ and $\beta = 0.5$. Find.

i) the probability that such a battery lasts more than 300 hours.

ii) mean ~~leaf~~ lifetime of these batteries.

X: lifetime (in hrs).

$X \sim$ Weibull $(\alpha = 0.1, \beta = 0.5)$.

i) $P[X > 300] = e^{-0.1(300)^{0.5}}$

$= e^{-0.1 \times 10 \times \sqrt{3}}$

$= e^{-\sqrt{3}} = 0.176$

ii) mean lifetime $= E(X)$

$= \frac{1}{\alpha^{1/\beta}} \Gamma\left(\frac{1}{\beta}+1\right)$

$= \frac{1}{(0.1)^2} \Gamma(2+1)$

$= \frac{1}{0.01} 2! = 200$ hrs

2. The lifetime (in hrs) of a component follows Weibull distribution with $\beta = 2$. From past experience, it is known that 15% of components fail before 100 hrs (the components that have lasted 90 hrs). Determine the value of α.

X : lifetime (in hrs).

$X \sim$ Weibull ($\alpha = ?$, $\beta = 2$).

$$P[X<100 \mid X>90] = 0.15.$$

$$\frac{P[X<100, X>90]}{P[X>90]} = 0.15$$

$$\frac{P[90<X<100]}{P[X>90]} = 0.15$$

$$\frac{F(100) - F(90)}{e^{-\alpha(90)^2}} = 0.15.$$

$$\frac{1 - e^{-\alpha(100)^2} - (1 - e^{-\alpha 90^2})}{e^{-\alpha(90)^2}}$$

$$-e^{-\alpha 100^2} + e^{-\alpha 90^2} = 0.15 e^{-\alpha 90^2}$$

$$-e^{-\alpha(100^2 - 90^2)} + 1 = 0.15$$

$$e^{-\alpha(1900)} = 0.85$$

$$\therefore -\alpha(1900) = \ln(0.85)$$

$$= -0.1625$$

$$\therefore \alpha = \frac{-0.1625}{-1900}$$

$$\boxed{\alpha = 0.0000855}$$

3. The lifetime of a certain component has Weibull distribution with $\beta = 2$. Find the value of α given that the probability that the component's life exceeds 5 yrs is $e^{-0.25}$. Find mean, variance.

X : lifetime (in yrs)

$X \sim$ Weibull $(\alpha = ?, \beta = 2)$.

$$P[X > 5] = e^{-0.25}$$

$$e^{-\alpha(5)^2} = e^{-0.25}.$$

$$25\alpha = 0.25.$$

$$\alpha = \frac{0.25}{25}.$$

$$\alpha = 0.01.$$

$$E(X) = \frac{1}{\alpha^{1/\beta}} \Gamma\left(\frac{1}{\beta}+1\right).$$

$$= \frac{1}{(0.01)^{1/2}} \Gamma\left(\frac{1}{2}+1\right).$$

$$= \frac{1}{0.1} \Gamma\left(\frac{1}{2}+1\right). \qquad \Gamma(n+1) = n\Gamma(n)$$

$$= 10 \, \frac{1}{2} \Gamma\left(\frac{1}{2}\right).$$

$$= 5\sqrt{\pi}.$$

$$E(X^2) = \frac{1}{\alpha^{2/\beta}} \Gamma\left(\frac{2}{\beta}+1\right)$$

$$= \frac{1}{(0.01)^{2/2}} \Gamma(2).$$

$$= 100 \times 1!$$

$$= 100.$$

$$Var(X) = 100 - 25\pi$$

Normal distribution / Gaussian distribution.

pdf : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$.

parameters : μ : mean

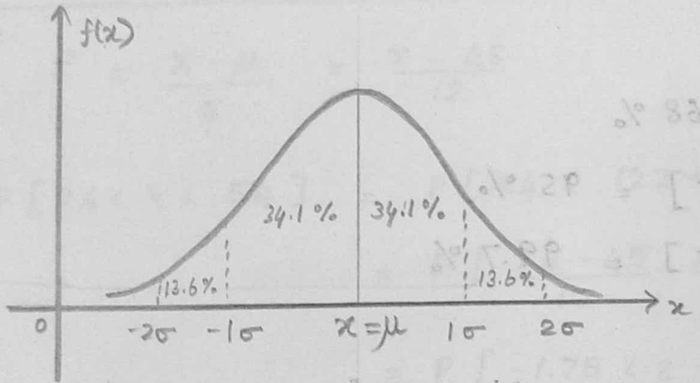
σ : standard deviation.

special case :

Standard normal distribution :

It is a ND with $\mu = 0$ and $\sigma = 1$

Note : graph of pdf of ND : Normal curve.



properties of normal curve :

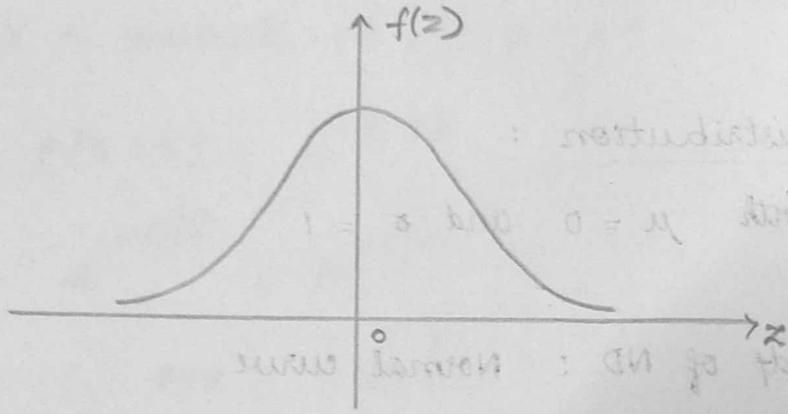
1. symmetric about the mean = μ.
2. mean = mode = median
3. $x = \mu - \sigma$, $x = \mu + \sigma$ are points of inflexion.
4. x-axis is the asymptote.
5. area under graph is 1.

Normal Table :

Table containing probabilities for standard Normal distribution. (SND)

An arbitrary Normal distribution with any mean μ and any standard deviation σ can be converted into a SND by the transformation

$Z = \frac{X - \mu}{\sigma}$ where Z is the standard Normal random variable.



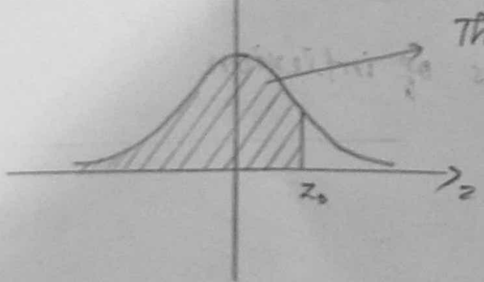
Empirical rule

$P[\mu - \sigma < X < \mu + \sigma] = 68\%$

$P[\mu - 2\sigma < X < \mu + 2\sigma] = 95\%$

$P[\mu - 3\sigma < X < \mu + 3\sigma] = 99.7\%$

From normal table one can find $P[Z \leq z_0] = \phi(z_0)$ for any $z_0 \in R$



Problem :

1. A survey indicates that for each trip to the super market a customer spends an average of 45 min with a S.D of 12 minutes. The length of time spent follows Normal distribution. What is the probability a randomly selected customer will spend.

i) b/w 24 and 54 minutes

ii) more than 39 minutes.

iii) b/w 33 and 60 minutes

X : time spent in the store (in minutes)

$X \sim ND(\mu = 45, \sigma = 12)$

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 45}{12}.$$

i) $P[24 < X < 54] = P[24 < 12Z + 45 < 54]$

$$= P\left[\frac{24-45}{12} < Z < \frac{54-45}{12}\right]$$

$$= P[-1.75 < Z < 0.75]$$

$$= \phi(0.75) - \phi(-1.75)$$

$$= 0.7734 - 0.0401$$

$$= 0.7333$$

ii) $P[X > 39] = P[12Z + 45 > 39]$

$$= P\left[Z > \frac{39-45}{12}\right]$$

$$= P[Z > -0.5]$$

$$= 1 - P[Z < -0.5]$$

$$= 1 - 0.3085$$

$$= 0.6915$$

iii) $P[33 < X < 60] = P[33 < 12z + 45 < 60]$

$= P\left[\frac{33-45}{12} < z < \frac{60-45}{12}\right]$

$= P[-1 < z < 1.25]$.

$= \phi(1.25) - \phi(-1)$

$= 0.8944 - 0.1587$

$= 0.7357$

2. The scores on a test given to 1 lakh students follows Normal distribution with a mean of 500 an S.D 100. What should be the score of student to place him among the top 10% ?

X : score on a test.

$X \sim ND(\mu = 500, SD = 100)$

$z = \frac{X - 500}{100} \Rightarrow X = 100z + 500$.

To find 'x' such that

$P[X \geq x] = 0.1$.

$P[100z + 500 \geq x] = 0.1$

$P\left[z \geq \frac{x-500}{100}\right] = 0.1$.

$\therefore P\left[z < \frac{x-500}{100}\right] = 0.9$

$\phi\left(\frac{x-500}{100}\right) = 0.9$

From table $\phi(1.28) = 0.8997 \simeq 0.9$

$\therefore \frac{x-500}{100} = 1.28$.

$x - 500 = 128$

$\boxed{x = 628.}$

3. In a normal distribution 31 % of the items are under 45 and 8 % of items are over 64. Find the mean and variance.

$X \sim ND(\mu, \sigma)$.

$P[X < 45] = 31\% = 0.31$

$P\left[Z < \frac{45-\mu}{\sigma}\right] = 0.31$.

$\phi\left(\frac{45-\mu}{\sigma}\right) = 0.31$

From table $\phi(-0.5) = 0.3085 \simeq 0.31$

$\therefore \frac{45-\mu}{\sigma} = -0.5$.

$\boxed{45 - \mu + 0.5\sigma = 0} \rightarrow ①$

$P[X > 64] = 0.08$.

$P\left[Z > \frac{64-\mu}{\sigma}\right] = 0.08$

$P\left[Z < \frac{64-\mu}{\sigma}\right] = 0.92$.

$\phi\left(\frac{64-\mu}{\sigma}\right) = 0.92$.

From table $\phi(1.41) = 0.9207 \simeq 0.92$

$\therefore \frac{64-\mu}{\sigma} = 1.41$

$\boxed{64 - \mu - 1.41\sigma = 0} \longrightarrow ②$

solving ① and ② :

$-19 + 1.91\sigma = 0$.

$1.91\sigma = 19$.

$\sigma = 9.95 \simeq 10$.

$\mu = 45 + 0.5\sigma = 49.98 \simeq 50$.

Moments (moment ... mean $E(X) = \mu$

Let X be a random variable with

$E(X^r)$ is called the r^{th} moment of X (r^{th} moment of X about the origin).

$E(X-\mu)^r$ is called the r^{th} central moment of X (r^{th} moment of X about the mean)

$r=1$

$E(X)$: First moment is the mean.

$r=2$

$E(X-\mu)^2$: Second central moment is the variance.

For other values of r also, $E(X^r)$ is a valuable measu...

NOTE :

$E(X-\mu)^r = \mu_r$

$E(X^r) = \mu'_r$.

The quantity $\frac{\mu_3}{\sigma^3}$ is a measure of symmetry / skewn...

$\frac{\mu_3}{\sigma^3} = 0 \Rightarrow$ symmetric. 

$\frac{\mu_3}{\sigma^3} > 0 \Rightarrow$ right skewed. 

$\frac{\mu_3}{\sigma^3} < 0 \Rightarrow$ left skewed. 

$$\frac{\mu_3}{\sigma^3} = \frac{E(X-\mu)^3}{\left[E(X-\mu)^2\right]^{3/2}}$$

The quantity $\frac{\mu_4}{\sigma^4}$ is a measure of Kurtosis (Peakne...

$\frac{\mu_4}{\sigma^4} = 3$ for standard Normal Distribution (SND)

$\frac{\mu_4}{\sigma^4} > 3$ more peaked than SND

$\frac{\mu_4}{\sigma^4} < 3$ more flat / less peaked than SND

MGF - Moment Generating Function

For a random variable X

$$\boxed{M_X(t) = E(e^{tX})}$$

If $M_X(t)$ is defined for all values of t in an interval $(-\delta, \delta)$ for $\delta > 0$, then $M_X(t)$ is called the moment generating function of X.

NOTE : $M_X(t)$ is finite in some neighbourhood of 0 i.e. $(-\delta, \delta)$. If the condition is not satisfied, some moments may not exist.

Moments of a r.v can be obtained from MGF

Finding / "generating" moments

Method 1 :

$$\frac{d^r}{dt^r}[M_X(t)]_{t=0} = E(X^r)$$

the rth derivative of MGF about $t=0$ is the rth moment.

Proof / verification.

$$M'_X(t) = \frac{d}{dt}\{E(e^{tX})\}$$

$$= E\left\{\frac{d}{dt}(e^{tX})\right\}$$

$$= E[Xe^{tX}]$$

$$M'_X(t)_{t=0} = E[X]$$

NOTE that first derivative gave us the first moment.

Method 2 :

$$M_x(t) = E(e^{tX}).$$

$$= E\left[1 + \frac{tX}{1!} + \frac{t^2X^2}{2!} + \dots + \frac{t^r X^r}{r!} + \dots\right]$$

$$= 1 + \frac{t}{1!}E(X) + \frac{t^2}{2!}E(X^2) + \dots + \frac{t^r}{r!}E(X^r) + \dots$$

The coefficient of $\frac{t^r}{r!}$ is the r^{th} moment

Problems :

1. Find the binomial distribution mgf given the parameters for BD are 'n' and 'p'. Hence find mean and variance.

pmf of BD : $f(x) = {}^nC_x p^x q^{n-x}, \quad x = 0, \dots n.$

mgf $\quad M_X(t) = E(e^{tX})$

$$= \sum e^{tx} f(x)$$

$$= \sum e^{tx}\, {}^nC_x p^x q^{n-x}$$

$$= \sum {}^nC_x (pe^t)^x q^{n-x}.$$

$$= (pe^t + q)^n.$$

First moment of X $= M'_X(t)\Big|_{t=0}$

$$E(X) = \left[n(pe^t+q)^{n-1} pe^t\right]_{t=0}$$

$$= n(p+q)^{n-1} p(1)$$

$$= n(1)^{n-1} p$$

$$\boxed{\text{mean} = np}$$

second moment of X $= M''_X(t)\Big]_{t=0}$

$$= \frac{d}{dt}\left[npe^t(pe^t+q)^{n-1}\right]_{t=0}$$

$$= np\left\{(pe^t+q)^{n-1}e^t + e^t(n-1).(pe^t+q)^{n-2}.pe^t\right\}_{t=0}$$

$$= np\{1+(n-1)p\}$$

$$= np\{1+np-p\}$$

$$= np\{q+np\}.$$

$$E(X^2) = n^2p^2 + npq.$$

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

$$= \cancel{n^2p^2} + npq - \cancel{n^2p^2}$$

$$\boxed{var(X) = npq.}$$

2. Find the mgf of exponential distribution and hence find mean, variance.

pdf of ED : $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$, $\lambda > 0$.

mgf $M_X(t) = E(e^{tX})$

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

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

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

$$= \lambda\left[0 - \frac{1}{t-\lambda}\right] \quad (\text{if } t<\lambda)$$

$$= \frac{-\lambda}{t-\lambda}$$

$$M_X(t) = \frac{\lambda}{\lambda-t}$$

First moment of x = $M_x'(t)_{t=0}$.

$$E(x) = \frac{-\lambda(-1)}{(\lambda-t)^2}_{t=0}$$

$$= \frac{\lambda}{(\lambda-t)^2}_{t=0}.$$

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

$$\boxed{\text{mean} = \frac{1}{\lambda}}$$

Second moment of x = $M_x''(t)_{t=0}$.

$$= \frac{d}{dt}\left[\frac{\lambda}{(\lambda-t)^2}\right]_{t=0}$$

$$= \frac{-\lambda(2)(\lambda-t)(-1)}{(\lambda-t)^4}_{t=0}$$

$$= \frac{2\lambda}{(\lambda-t)^3}_{t=0}.$$

$$= \frac{2\lambda}{\lambda^3}$$

$$E(x^2) = \frac{2}{\lambda^2}$$

$$var(x) = E(x^2) - E(x)^2$$

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

$$\boxed{var(x) = \frac{1}{\lambda^2}}$$

3. Find the mgf of Poisson distribution and hence find its mean and variance.

pmf of PD : $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0, 1, 2, \ldots$ $\lambda > 0$

mgf $M_x(t) = E(e^{tx})$

$$= \sum e^{tx} \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \sum \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^t}$$

$$= e^{\lambda(e^t - 1)}$$

First moment of $x = M'_x(t)_{t=0}$.

$$E(x) = e^{\lambda(e^t-1)} \lambda e^t \Big|_{t=0}.$$

$$= 1(\lambda)$$

$$\boxed{\text{mean} = \lambda.}$$

second moment of $x = M''_x(t)_{t=0}$.

$$E(x^2) = e^{\lambda(e^t-1)} \lambda e^t + \lambda e^t e^{\lambda(e^t-1)} \lambda e^t \Big|_{t=0.}$$

$$= \lambda + \lambda^2.$$

$\text{var}(x) = E(x^2) - E(x)^2.$

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

$$\boxed{\text{var}(x) = \lambda.}$$

4. Find the mgf of standard Normal Distribution and hence find mean and variance.

$z \sim$ SND.

pdf of SND : $g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad -\infty < z < \infty$

mgf $M_z(t) = E(e^{tz})$

$$= \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{(tz - z^2/2)} dz.$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z^2 - 2tz)} dz.$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z^2 - 2tz + t^2 - t^2)} dz.$$

$$= \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z-t)^2} dz.$$

Let $u = z - t$.

$du = dz$

$$= \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du.$$

$$= e^{t^2/2} (1) \quad \because \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} \text{ is the pdf}$$

the standard normal 'u'.

$$\boxed{M_z(t) = e^{t^2/2}}$$

First moment of z $= M_z'(t)_{t=0}$.

$$E(z) = e^{\frac{t^2}{2}} \frac{\cancel{2}t}{\cancel{2}} \Big|_{t=0}.$$

$$= te^{t^2/2}. \Big|_{t=0}$$

$$\boxed{\text{mean} = 0.}$$

second moment of z $= M_z''(t)_{t=0}$.

$$= e^{t^2/2} + t^2 e^{t^2/2} \Big|_{t=0}.$$

$= e^0 + 0$

$E(z^2) = 1$

$var(z) = E(z^2) - E(z)^2$

$= 1 - 0$

$\boxed{var(z) = 1.}$

Properties of mgf :

Two random variables x and Y with same mgf have the same probability distribution. i.e. mgf of a random variable - unique

$M_x(0) = 1$

$M_{cx}(t) = M_x(ct)$ [Property (i)]

$M_{c+x}(t) = e^{ct} M_x(t)$ [Property (ii)]

mgf of 2 independent random variables' sum

$M_{x+y}(t) = M_x(t) . M_y(t)$

5. Find the mgf of normal distribution from that of standard normal distribution.

mgf of z : $M_z(t) = e^{t^2/2}$

and $z = \frac{x-\mu}{\sigma} \Rightarrow x = \mu + \sigma z$.

mgf of Normal variable :

$M_x(t) = M_{\mu+\sigma z}(t)$

$= e^{\mu t} M_{\sigma z}(t)$ using (ii)

$= e^{\mu t} M_z(\sigma t)$ using (i)

$$= e^{\mu t} e^{\frac{(\sigma t)^2}{2}}$$

$$\boxed{M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}}$$

6. Use the mgf of $ND(\mu, \sigma)$ to find mean and variance.

$$\text{mgf} \quad M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

First moment of x $= M'_X(t) \big|_{t=0}$.

$$E(x) = e^{\mu t + \frac{\sigma^2 t^2}{2}} (\mu + \sigma^2 t) \Big|_{t=0}$$

$$= e^0 (\mu + 0)$$

$$= 1(\mu)$$

$$\boxed{\text{mean} = \mu.}$$

Second moment of x $= M''_X(t) \big|_{t=0}$.

$$E(x^2) = e^{\mu t + \frac{\sigma^2 t^2}{2}} (\mu + \sigma^2 t)^2 + e^{\mu t + \frac{\sigma^2 t^2}{2}} \sigma^2 \Big|_{t=0}$$

$$= \mu^2 + \sigma^2.$$

$$var(x) = E(x^2) - E(x)^2$$

$$= \cancel{\mu^2} + \sigma^2 - \cancel{\mu^2}$$

$$\boxed{var(x) = \sigma^2}$$

7. Let x be a random variable with mgf $M_X(t) = e^{2t^2}$. Find $P(0 < x < 1)$

comparing e^{2t^2} with $e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$$\mu = 0 \quad \text{and} \quad \frac{\sigma^2}{2} = 2 \quad \Rightarrow \sigma = 2.$$

$$x \sim N(\mu = 0, \sigma = 2).$$

$$z = \frac{x-\mu}{\sigma} = \frac{x}{2}.$$

$$p(0 < x < 1) = P(0 < z < 0.5)$$

$$= \phi(0.5) - \phi(0)$$

$$= 0.6915 - 0.5$$

$$= 0.1915$$

8. Let x be a random variable with pmf $p(x) = 2\left(\frac{1}{3}\right)^x$, $x = 1, 2, \dots$ Find the mgf of x and find mean.

mgf, $M_x(t) = E(e^{tx})$

$$= \sum e^{tx}\, 2\left(\frac{1}{3}\right)^x, \quad x = 1, 2, \dots$$

$$= 2 \sum \left(\frac{e^t}{3}\right)^x$$

$$= 2\left\{ \frac{e^t}{3} + \left(\frac{e^t}{3}\right)^2 + \dots \right\}$$

$$= 2\left[\frac{1}{1-\frac{e^t}{3}} - 1\right]$$

$$= 2\left[\frac{3}{3-e^t} - 1\right]$$

$$\boxed{M_x(t) = \frac{6}{3-e^t} - 2.}$$

First moment of x $= M'_x(t)\big|_{t=0}$.

$$E(X) = \frac{-6(-1)e^t}{(3-e^t)^2}\Bigg|_{t=0}.$$

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

$$\boxed{\text{mean} = 3/2}$$

Sum of independent random variables

If $x_1, x_2, \dots x_n$ are independent random variables with mgfs $M_{x_1}(t), M_{x_2}(t), \dots, M_{x_n}(t)$, then the mgf of $x = x_1 + x_2 + x_3 + \dots + x_n$,

$$\boxed{M_x(t) = M_{x_1}(t) \times M_{x_2}(t) \times \dots \times M_{x_n}(t)}$$

Result:

i) Sum of 2 independent normal random variables is a normal random variable.

proof:

$X_1 \sim N(\mu_1, \sigma_1)$

$X_2 \sim N(\mu_2, \sigma_2)$.

Let $X = X_1 + X_2$.

$$M_X(t) = M_{X_1}(t) \times M_{X_2}(t)$$

$$= e^{\mu_1 t + \frac{\sigma_1^2 t^2}{2}} \cdot e^{\mu_2 t + \frac{\sigma_2^2 t^2}{2}}$$

$$= e^{(\mu_1+\mu_2)t + \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}}$$

$\therefore$ X is a normal random variable with mean '$\mu_1+\mu_2$' and variance '$\sigma_1^2 + \sigma_2^2$'.

ii) If $X_1, X_2, \dots X_k$ are binomial random variables with parameters $(n_1, p), (n_2, p), \dots, (n_k, p)$ that are independent then $X_1 + X_2 + \dots + X_k$ is a binomial random variable with parameters $(n_1 + n_2 + \dots + n_k, p)$.

proof:

$$X = X_1 + X_2 + \dots + X_k.$$

$$M_X(t) = M_{X_1}(t) \times M_{X_2}(t) \times \dots \times M_{X_k}(t)$$

$$= (pe^t + q)^{n_1} \times (pe^t + q)^{n_2} \times \dots \times (pe^t + q)^{n_k}$$

$$= (pe^t + q)^{n_1 + n_2 + \dots + n_k}$$

iii) If $x_1, x_2, \ldots, x_n$ are independent poisson random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $x_1 + x_2 + \ldots x_n$ is a poisson random variable with mean $\lambda_1 + \lambda_2 + \ldots \lambda_n$.

proof :

$$x = x_1 + x_2 + \ldots + x_n.$$

$$M_x(t) = M_{x_1}(t) . M_{x_2}(t) \ldots M_{x_n}(t).$$

$$= e^{\lambda_1(e^t-1)} . e^{\lambda_2(e^t-1)} \ldots e^{\lambda_n(e^t-1)}.$$

$$= e^{(\lambda_1+\lambda_2+\ldots+\lambda_n)(e^t-1)}.$$

Erlang distribution :

generalisation of exponential distribution.

While exponential distribution describes the time b/w successive occurrences of an event, k-Erlang describes the time b/w any occurrence and the k^{th} occurrence of the event.

pdf : $f(x) = \dfrac{\lambda e^{-\lambda x} (\lambda x)^{k-1}}{\Gamma(k)}$, $x \geq 0$.

parameters : λ and k.

NOTE :

$k=1$ reduces Erlang distribution to Exponential distribution with parameter λ.

The generalised form of Erlang distribution is called Gamma distribution.

For any real no 'r', a continuous random variable x follows Gamma distribution if its pdf is given by

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{r-1}}{\Gamma(r)}, \quad x \geq 0$$

$X \sim$ Gamma (r, λ).

mean : $\frac{r}{\lambda}$

variance : $\frac{r}{\lambda^2}$

Mgf of r-Erlang distribution

pdf of r-Erlang : $f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{r-1}}{\Gamma(r)}$ $x \geq 0$.

mgf of X : $M_X(t) = E(e^{tX})$

$$= \int_0^\infty e^{tx} \frac{\lambda e^{-\lambda x}(\lambda x)^{r-1}}{\Gamma(r)} dx$$

$$= \frac{\lambda^r}{\Gamma(r)} \int_0^\infty e^{-(\lambda-t)x} x^{r-1} dx$$

$\int_0^\infty e^{-ax} x^n dx = \frac{n!}{a^{n+1}}$

$$= \frac{\lambda^r}{\cancel{\Gamma(r)}} \frac{\cancel{\Gamma(r)}}{(\lambda-t)^r}$$

$$= \frac{\lambda^r}{(\lambda-t)^r}$$

$$= \left(\frac{\lambda}{\lambda-t}\right)^r$$

Result (continued...)

iv) sum of r independent identical exponential random variables with parameter λ is an Erlang random variable with parameter r and λ.

proof :

$$X = X_1 + X_2 + \dots + X_r$$

$$M_X(t) = M_{X_1}(t) \cdot M_{X_2}(t) \dots M_{X_r}(t)$$

$$= \left(\frac{\lambda}{\lambda-t}\right)\left(\frac{\lambda}{\lambda-t}\right) \dots \left(\frac{\lambda}{\lambda-t}\right)$$

$$= \left(\frac{\lambda}{\lambda-t}\right)^r$$

Joint probability distribution

2 or more random variables defined on the same sample space. ie X & Y

joint pdf : $f_{XY}(x,y)$

joint pmf : $P_{XY}(x,y) = P[X=x, Y=y]$

joint cdf : $F_{XY}(x,y) = P[X \le x, Y \le y]$

The individual probability distribution of random variables X and Y are called marginal distributions denoted by $f_X(x)$, $f_Y(y)$

Marginal distributions of X and Y :

Let $P_{XY}(x,y)$ be the joint pmf of X and Y.

then

marginal distribution of X : $P_X(x) = \sum_{y \in Y} P_{XY}(x,y)$.

marginal distribution of Y: $P_Y(y) = \sum_{x \in X} P_{XY}(x,y)$

Example :

2 dice are rolled. Let X denote the maximum of the two throws, Y denote the number of times an even number appears. Find the joint pmf of X and Y. Also find the marginal distributions of X and Y.

$X = \{1,2,3,4,5,6\}$

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

X \ Y	0	1	2		
1	1/36	0	0	1/36	$P_X(1)$
2	0	2/36	1/36	3/36	$P_X(2)$
3	3/36	2/36	0	5/36	$P_X(3)$
4	0	4/36	3/36	7/36	$P_X(4)$
5	5/36	4/36	0	9/36	$P_X(5)$
6	0	6/36	5/36	11/36	$P_X(6)$
	9/36	18/36	9/36	1	
	$P_Y(0)$	$P_Y(1)$	$P_Y(2)$		

marginal distribution of X

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

marginal distribution of Y

Y	0	1	2
$P_Y(y)$	9/36	18/36	9/36

Properties of joint pmf / pdf :

pmf :

i) $P_{XY}(x,y) \geq 0$.

ii) $\sum_{x \in X} \sum_{y \in Y} P_{XY}(x,y) = 1$

pdf :

i) $f_{XY}(x,y) \geq 0$

ii) $\int\int f_{XY}(x,y)\,dy\,dx = 1$

If $f_{XY}(x,y)$ is the joint pdf of X and Y, then

marginal distribution of X : $f_X(x) = \int f_{XY}(x,y)\,dy$

marginal distribution of Y : $f_Y(y) = \int f_{XY}(x,y)\,dx$.

NOTE :

If X and Y are independent random variables then

$f_{XY}(x,y) = f_X(x) . f_Y(y)$.

conditional distribution :

For 2 events A, B.

$$P(A/B) = \frac{P(A \cap B)}{P(B)} , \quad P(B) \neq 0.$$

For 2 random variables X and Y with joint pmf/pdf $f(x,y)$, the conditional distribution of Y given X = x. is given by.

$$f_{Y|X=x}(y|x) = \frac{f(x,y)}{f_X(x)}, \quad f_X(x) \neq 0$$

Conditional expectation and variance :

The conditional expectation of Y given X = x. is given by

$$E[Y|X=x] = \sum_{y \in Y} y \, f_{Y|X=x}(y|x). \qquad \text{if } X, Y \text{ are discrete}.$$

$$E[Y|X=x] = \int y \, f_{Y|X}(y|x) \, dy \qquad \text{if } X, Y \text{ are continuous}$$

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

Result :

1. For 2 random variables X and Y with join pmf/pdf $f(x,y)$ $E[E(X|Y)] = E(X)$.

proof :

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

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

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

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x,y)\, dy \right) x\, dx$$

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

$$= E[X]$$

Problems:

1. The joint pdf of r.vs X and Y is $f(x,y) = \begin{cases} kxy^2, & 0 \le x \le \\ 0, & \text{other} \end{cases}$

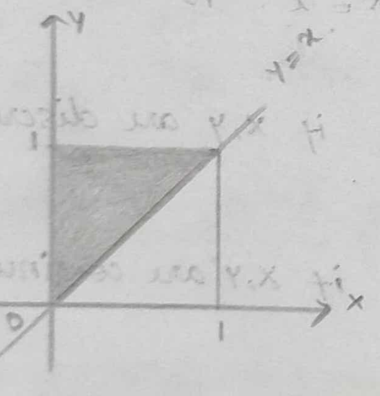
Find i) the value of k.

ii) marginal distributions of X and Y

iii) $E(X)$ and $E(Y)$

iv) if X and Y are independent.

i)



$$\iint f(x,y)\, dy\, dx = 1$$

$$\int_0^1 \int_x^1 kxy^2\, dy\, dx = 1.$$

$$\int_0^1 \left[\frac{kxy^3}{3} \right]_x^1 dx = 1.$$

$$\int_0^1 \left(\frac{kx}{3} - \frac{kx^4}{3} \right) dx = 1.$$

$$\frac{k}{3} \left[\frac{x^2}{2} - \frac{x^5}{5} \right]_0^1 = 1$$

$$\frac{k}{3}\left[\frac{1}{2}-\frac{1}{5}-0\right] = 1.$$

$$k = 3 \times \frac{10}{3}$$

$$\boxed{k = 10}$$

ii)

$$f_X(x) = \int_x^1 f(x,y)\,dy.$$

$$= \int_x^1 10xy^2\,dy.$$

$$= 10x\left[\frac{y^3}{3}\right]_x^1.$$

$$= 10x\left(\frac{1}{3}-\frac{x^3}{3}\right).$$

$$f_X(x) = \frac{10(x-x^4)}{3.}, \quad 0 \le x \le 1.$$

$$f_Y(y) = \int f(x,y)\,dx.$$

$$= \int_0^y 10xy^2\,dx.$$

$$= \frac{10}{2}y^2\left[x^2\right]_0^y$$

$$= 5y^2(y^2)$$

$$f_Y(y) = 5y^4. \quad , \quad 0 \le y \le 1$$

iii) $E(X) = \int x f_X(x)\,dx$

$$= \frac{10}{3}\int_0^1 (x^2 - x^5)\,dx$$

$$= \frac{10}{3}\left[\frac{x^3}{3}-\frac{x^6}{6}\right]_0^1$$

$$= \frac{10}{3}\left[\frac{1}{3}-\frac{1}{6}-0\right].$$

$$= \frac{\cancel{10}^{5}}{3}\,\frac{1}{\cancel{6}_{3}}$$

$$= \frac{5}{9}$$

$E(Y) = \int y f_Y(y)\,dy$

$$= \int_0^1 5y^5\,dy.$$

$$= \frac{5}{6}\left[y^6\right]_0^1$$

$$= \frac{5}{6}(1-0)$$

$$= \frac{5}{6}$$

iv) $f_X(x) . f_Y(y) = \frac{10}{3} x(1-x^3) . 5y^4.$

$$\neq f_{XY}(x,y).$$

$\therefore$ X and Y are <u>dependent</u>.

<u>Aliter :</u>

$$E(XY) = \iint xy\, f(x,y)\, dy\, dx.$$

$$= \int_0^1 \int_x^1 xy\, 10xy^2\, dy\, dx.$$

$$= \int_0^1 10x^2 \left[\frac{y^4}{4}\right]_x^1 dx.$$

$$= 10 \int_0^1 x^2 \left[\frac{1}{4} - \frac{x^4}{4}\right] dx.$$

$$= \frac{10}{4} \int_0^1 (x^2 - x^6)\, dx.$$

$$= \frac{10}{4} \left[\frac{x^3}{3} - \frac{x^7}{7}\right]_0^1$$

$$= \frac{10}{4} \left[\frac{1}{3} - \frac{1}{7}\right]$$

$$= \frac{10}{\cancel{4}} \left[\frac{\cancel{4}}{21}\right]$$

$$= \frac{10}{21}$$

$$E(X). E(Y) = \frac{5}{9} \cdot \frac{5}{6}$$

$$= \frac{25}{54}.$$

$$E(XY) \neq E(X). E(Y).$$

$\therefore$ X and Y are <u>dependent</u>

2. $f(x,y) = \begin{cases} xy^2 + \frac{x^2}{8} & , 0 \le x \le 2, \quad 0 \le y \le 1 \\ 0 & , \text{otherwise} \end{cases}$

Find i) $P[X>1]$

ii) $P[Y < 1/2]$

iii) $P[X>1, Y<1/2]$

iv) $P[X<Y]$

v) $P[Y < 1/2 \mid X > 1]$

vi) $P[X+Y \le 1]$

$f_X(x) = \int f(x,y)\,dy$.

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

$= \left[x\frac{y^3}{3} + \frac{x^2}{8}y\right]_0^1$

$= \frac{x}{3} + \frac{x^2}{8} - 0$.

$= \frac{x}{3} + \frac{x^2}{8}, \quad 0 \le x \le 2$

$f_Y(y) = \int f(x,y)\,dx$.

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

$= \left[y^2\frac{x^2}{2} + \frac{x^3}{24}\right]_0^2$.

$= 2y^2 + \frac{8}{24} - 0$

$= 2y^2 + \frac{1}{3}, \quad 0 \le y \le 1$.

i) $P[X>1] = \int_1^\infty f_X(x)\,dx$.

$= \int_1^2 \left[\frac{x}{3} + \frac{x^2}{8}\right] dx$.

$= \left[\frac{x^2}{6} + \frac{x^3}{24}\right]_1^2$

$= \frac{4}{6} + \frac{8}{24} - \frac{1}{6} - \frac{1}{24}$

$= \frac{3}{6} + \frac{7}{24} = \frac{19}{24}$

ii) $P[Y < 1/2] = \int_{-\infty}^{1/2} f_Y(y)\,dy$.

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

$= \left[\frac{2}{3}y^3 + \frac{y}{3}\right]_0^{1/2}$

$= \frac{\cancel{2}}{3} \times \frac{1}{\cancel{8}4} + \frac{1}{6} - 0.$

$= \frac{1}{12} + \frac{1}{6}.$

$= \frac{3}{12}$

$= \frac{1}{4}$

iii) $P[X > 1, Y < 1/2] = \int_1^{\infty}\int_{-\infty}^{1/2} f(x,y)\,dy\,dx$.

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

$= \int_1^2 \left\{x\left[\frac{y^3}{3}\right]_0^{1/2} + \frac{x^2}{8}[y]_0^{1/2}\right\} dx.$

$= \int_1^2 \left(\frac{x}{24} + \frac{x^2}{16}\right) dx.$

$= \frac{1}{24}\left[\frac{x^2}{2}\right]_1^2 + \frac{1}{16}\left[\frac{x^3}{3}\right]_1^2$

$= \frac{1}{24}\left[\frac{3}{2}\right] + \frac{1}{16}\left[\frac{7}{3}\right]$

$= 3/48 + 7/48$

$= 10/48$

$= 5/24$

iv) $P(X < Y) = \int\int f(x,y)\,dy\,dx$;

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

$$= \int_0^1 \left(\frac{x}{3}[y^3]_x^1 + \frac{x^2}{8}[y]_x^1 \right) dx$$

$$= \int_0^1 \left[\frac{x}{3}(1-x^3) + \frac{x^2}{8}(1-x) \right] dx$$

$$= \int_0^1 \left[\frac{x}{3} - \frac{x^4}{3} + \frac{x^2}{8} - \frac{x^3}{8} \right] dx$$

$$= \left[\frac{x^2}{6} - \frac{x^5}{15} + \frac{x^3}{24} - \frac{x^4}{32} \right]_0^1$$

$$= \frac{1}{6} - \frac{1}{15} + \frac{1}{24} - \frac{1}{32}$$

$$= \frac{4+1}{24} - \frac{1}{15} - \frac{1}{32}$$

$$= \frac{5}{24} - \frac{1}{15} - \frac{1}{32}$$

$$= \frac{100-32-15}{480}$$

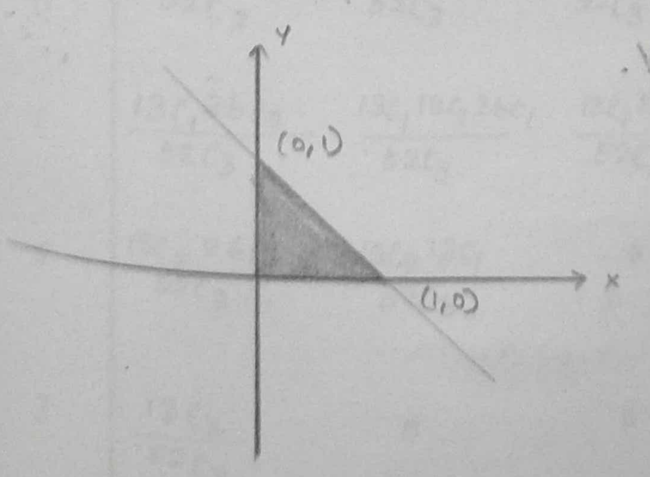
$$= \frac{53}{480}$$

v) $P[Y < 1/2 \mid X > 1] = \dfrac{P[X>1, Y<1/2]}{P[X>1]}$

$$= \frac{5/24}{19/24}$$

$$= \frac{5}{19}$$

vi) $P[X+Y \le 1]$



$P[X+Y \le 1] = \iint f(x,y)\,dy\,dx.$

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

$= \int_0^1 \left(\frac{x}{3}[y^3]_0^{1-x} + \frac{x^2}{8}[y]_0^{1-x}\right) dx.$

$= \int_0^1 \left(\frac{x}{3}(1-x)^3 + \frac{x^2}{8}(1-x)\right) dx.$

$= \int_0^1 \frac{1}{3}(x - 3x^2 + 3x^3 - x^4) + \frac{1}{8}(x^2 - x^3)\,dx.$

$= \frac{1}{3}\left[\frac{x^2}{2} - x^3 + \frac{3}{4}x^4 - \frac{x^5}{5}\right]_0^1 + \frac{1}{8}\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1$

$= \frac{1}{3}\left[\frac{1}{2} - 1 + \frac{3}{4} - \frac{1}{5}\right] + \frac{1}{8}\left[\frac{1}{3} - \frac{1}{4}\right]$

$= \frac{1}{3}\left[\frac{10-20+15-4}{20}\right] + \frac{1}{8}\left[\frac{1}{12}\right].$

$= \frac{1}{60} + \frac{1}{96}.$

$= \frac{8+5}{480}$

$= 13/480.$

3. Roll a fair die. Let the outcome be X. Then toss a fair coin X times and let Y denote the no. of tails. Find the joint pmf of X and Y. Also find the marginal pmf of X and Y.

$X = \{1, 2, 3, 4, 5, 6\}$

$Y = \{0, 1, 2, 3, 4, 5, 6\}$

X \ Y	0	1	2	3	4	5	6
1	1/12	1/12	0	0	0	0	0
2	1/24	2/24	1/24	0	0	0	0
3	1/48	3/48	3/48	1/48	0	0	0
4	1/96	4/96	6/96	4/96	1/96	0	0
5	1/192	5/192	10/192	10/192	5/192	1/192	0
6	1/384	6/384	15/384	20/384	15/384	6/384	1/384

Note that $p(x,y) = \frac{1}{6}\, {}^{x}C_y \left(\frac{1}{2}\right)^{x}$, $y = 0,1,2,\ldots x$.

marginal distribution of X:

$P_X(x) = \frac{1}{6}$, $x = 1,2,3,4,5,6$.

marginal distribution of Y:

$Y=y$	0	1	2	3	4	5	6
$P_Y(y)$	$\frac{63}{384}$	$\frac{120}{384}$	$\frac{99}{384}$	$\frac{64}{384}$	$\frac{29}{384}$	$\frac{8}{384}$	$\frac{1}{384}$.

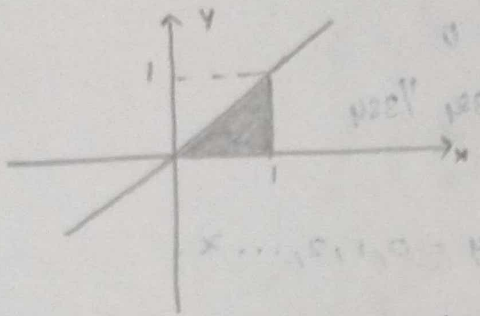
4. suppose 3 cards are drawn from a deck of 52 cards if X and Y denote the no. of diamonds and spades respectively. Find the joint pmf of X & Y.

X \ Y	0	1	2	3
0	$\frac{26C_3}{52C_3}$	$\frac{26C_2\, 13C_1}{52C_3}$	$\frac{26C_1\, 13C_2}{52C_3}$	$\frac{13C_3}{52C_3}$
1	$\frac{13C_1\, 26C_2}{52C_3}$	$\frac{13C_1\, 13C_1\, 26C_1}{52C_3}$	$\frac{13C_1\, 13C_2}{52C_3}$	0
2	$\frac{13C_2\, 26C_1}{52C_3}$	$\frac{13C_2\, 13C_1}{52C_3}$	0	0
3	$\frac{13C_3}{52C_3}$	0	0	0

5. The joint pdf of x and y is

$$f(x,y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \quad 0<x<1,\ 0<y<2.$$ Find

i) $P[X>Y]$.



$$P[X>Y] = \int\int f(x,y)\,dy\,dx.$$

$$= \frac{6}{7}\int_0^1\int_0^x \left(x^2+\frac{xy}{2}\right)dy\,dx.$$

$$= \frac{6}{7}\int_0^1 \left(x^2[y]_0^x + \frac{x}{2(2)}[y^2]_0^x\right)dx.$$

$$= \frac{6}{7}\int_0^1 \left(x^3 + \frac{x^3}{4}\right)dx.$$

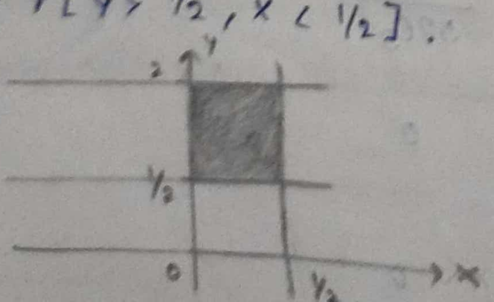
$$= \frac{6}{7}\left[\frac{x^4}{4} + \frac{x^4}{16}\right]_0^1$$

$$= \frac{6}{7}\left[\frac{1}{4}+\frac{1}{16}-0\right].$$

$$= \frac{\cancel{6}\,3}{7}\left[\frac{5}{\cancel{16}\,8}\right].$$

$$= \frac{15}{56}$$

ii) $P[Y>1/2, X<1/2]$.



$$P[Y>1/2, X<1/2] = \int_{-\infty}^{1/2}\int_{1/2}^{\infty} f(x,y)\,dy\,dx.$$

$$= \frac{6}{7} \int_0^{1/2} \int_{1/2}^{2} \left(x^2 + \frac{xy}{2} \right) dy\, dx.$$

$$= \frac{6}{7} \int_0^{1/2} \left(x^2 [y]_{1/2}^{2} + \frac{x}{4} [y^2]_{1/2}^{2} \right) dx.$$

$$= \frac{6}{7} \int_0^{1/2} \left(\frac{3}{2} x^2 + \frac{15}{16} x \right) dx.$$

$$= \frac{6}{7} \left(\frac{1}{2} [x^3]_0^{1/2} + \frac{15}{32} [x^2]_0^{1/2} \right).$$

$$= \frac{6}{7} \left(\frac{1}{2} \frac{1}{8} + \frac{15}{32} \frac{1}{4} \right).$$

$$= \frac{6}{7} \left(\frac{8}{128} + \frac{15}{128} \right).$$

$$= \frac{\overset{3}{\cancel{6}}}{7} \left(\frac{23}{\underset{64}{\cancel{128}}} \right).$$

$$= \frac{23 \times 3}{448}$$

$$= \frac{69}{448}$$