* Random Experiment:-
  Experiment whose outcome is unknown

eg.. Toss a coin, roll a dice, time of waiting inqueue.

* sample space :-
  set of all possible outcomes.

for coin → $\Omega = [H, T]$. Ω - sample space.

for dice → $\Omega = [1, 2, 3, 4, 5, 6]$

for queue → $\Omega = [0, \infty)$.

sample space may be finite / infinite.

It may or may not be a subset of numberline.

* Event:- subset of sample space you are interested in.

sample space
- discrete ([H,T] [1,2,3,4,5,6] [N])
- continuous [[a,∞), [a,b]).

* Assignment of probability:-

• case 1: Discret finite & equally likely outcomes

Let A be an event then,

$$P(A) = \frac{N(A)}{N(\Omega)}$$ → relative freq^n approach.

• Case 2: Discret finite + All outcomes are not equally likely.

- Run the exp. N no. of times.
- By F_A denote the exhaustive Freq^n of A

$$\frac{(\text{No. of times } A \text{ occured in } n \text{ times})}{n}$$

$$= P(A) = \lim_{n \to \infty} F_A$$

Observation:- only for certain key events one needs to assign probability.

eg., for a die $\rightarrow$ we need to assign only $P(1), P(2) \dots P(5)$
Then all other events are automatically satisfied
$\downarrow$
How?

- we intuitively assume certain laws of probability.
These assumptions are formally formed into axioms.

- why axiomatic defination?

$\Rightarrow$ Works in all possible cases.

$\Rightarrow$ Generalizes the ideas of probability assignment.

Event $\rightarrow$ subset of Ω

$\cup$ & $\cap$ of events is also an event.

σ-field:- A collection of sets $\mathcal{F}$ is called σ-algebra iff.

i) $A \in \mathcal{F}$ then, $A^c \in \mathcal{F}$ [closed under complement]

ii) $A_1, A_2 \dots \in \mathcal{F}$ then,

$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ & $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$

countable union & intersection.

eg: $\Omega = [H, T]$

$\mathcal{F} = 2^{\Omega} = \{\phi, [H], [T], [H,T]\}$.

Power set of Ω.

(Power set $\rightarrow$ set of all subsets)

Power set is always σ-algebra.

eg., $\Omega = [1, 2, 3, 4, 5, 6]$

$\mathcal{F} = 2^{\Omega}$ (2^6 elements in $\mathcal{F}$).

Axiomatic Defination of probability:-

Let Ω be the sample space. Let the σ-algebra associated with Ω.

Then a probability measure P is a function of $\mathcal{F}$ to R

($P : \mathcal{F} \rightarrow R$) such that,

i) $P(A) \geq 0$, $\forall A \in \mathcal{F}$

ii) $P(\Omega) = 1$.

iii) If A_i are mutually disjoint for $i = 1, 2, \dots$

($A_i \cap A_j = \phi$ $\forall i \neq j$)

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

• Properties:-

i) $P(A^c) = 1 - P(A)$ [From (ii) & (iii)].

$A \cup A^c = \Omega$ $\therefore P(A \cup A^c) = P(A) + P(A^c)$

$P(\Omega) = P(A) + P(A^c) = 1$.

$\Rightarrow P(A^c) = 1 - P(A)$.

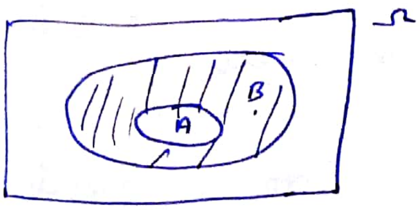
ii) $P(\phi) = 0$.

take $A = \Omega$ in previous result $\therefore A^c = \phi$.

$\therefore P(A^c) = 1 - P(\Omega) = 1 - 1 = 0$.

$\therefore P(\phi) = 0$.

iii) If $A \subset B$ then, $P(A) \leq P(B)$.
↳ A is subset of B.



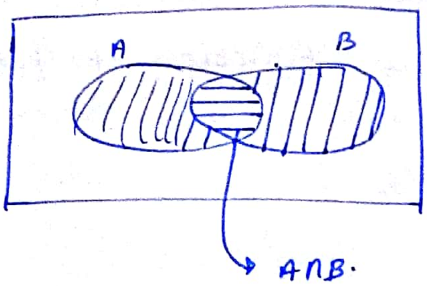
▨ - shaded part.

$B = A + (B \cap A^c)$

$\therefore P(B) = P(A) + P(B \cap A^c)$
≥ 0 by (i).

$\therefore P(A) \leq P(B)$

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



Ω $\quad A \cup B = \underbrace{A \cap B^c + B \cap A^c + A \cap B}_{\text{disjoint union}}$.

By 3rd axiom →

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

$A = A \cap B^c + A \cap B$.

$\therefore P(A) = P(A \cap B^c) + P(A \cap B)$.

$\therefore P(A \cap B^c) = P(A) - P(A \cap B)$

parallelly, $P(B \cap A^c) = P(B) - P(A \cap B)$.

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

[write down as a union of disjoint sets such that you can count the probability of disjoint sets]

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

(True by induction continuously)

eg., Toss 2 fair coins.

$\Omega = \{(HH), (HT), (TH), (TT)\}$.

$\mathcal{F} = 2^{\Omega}$

Assume → $P(HH) = 1/4$ $\quad P(TH) = 1/4$.

$P(HT) = 1/4$ $\quad P(TT) = 1/4$.

↳ valid probability assignment

all rules of axiomatic defination sets.

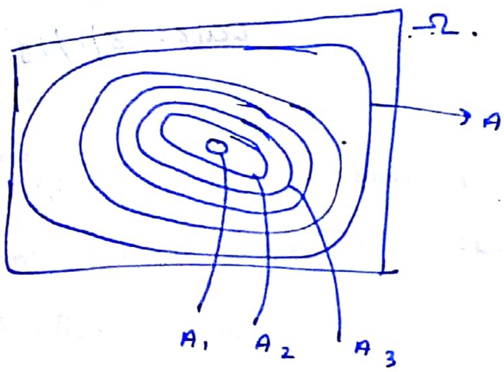
Theorem:- Let $A_1 \subset A_2 \subset A_3 \cdots A$ & $A = \bigcup_{i=1}^{n} A_i$

increasing sets then,

$P(A) = \lim_{n\to\infty} P(A_n)$

is subset of

eg : $0 \subset (-1,1) \subset (-2,2) \subset \cdots\cdots$



If $B_1 = A_1$, $B_2 = A_2 \cap A_1^c$, $B_3 = A_3 \cap A_2^c$ & so on

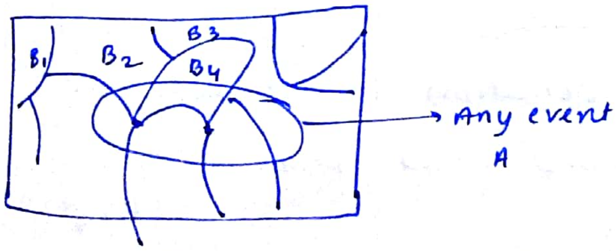
$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i$

$A = \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i$

$\therefore P(A) = P(\bigcup_{i=1}^{n} B_i) \to$ Probability of disjoint unions.

$P(A) = \sum_{i=1}^{n} P(B_i)$

Apply de-Morgan's to get result.



→ $B_1, B_2, \cdots B_n$ is a disjoint cover of Ω.

A intersects with B_i's & it can be written as a disjoint union of $A \cap B_i$'s

* Disjoint cover →

$$\left[\begin{array}{ll} B_i \cap B_j = \phi & \forall i \neq j. \\ \bigcup_{i=1}^{n} B_i = \Omega & \end{array}\right].$$

* $P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_n).$

---

* **Conditional Probability:-** Date: 8/1/19.

It modifies the probability.

What is the probability that event A occurs given event B has already occured:

$A, B \in \mathcal{F}$ s.t. $P(B) > 0.$

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

R. exp. → Rolling a die.

A → 2 appears on top face $\{2\}$.

B → even number appears $\{2, 4, 6\}$.

$$P(A|B) = \frac{\text{no. of } A}{\text{no. of } B} = \frac{1}{3}. = \frac{\# A}{\# B}$$

$$= \frac{\#(A \cap B)}{\# B} = \underbrace{\frac{\#(A \cap B) / \#(S.P.)}{\#(B) / \#(S.P.)}}_{\text{Motivation for formula of conditional prob.}}$$

Motivation for formula of conditional prob.

$\boxed{P(A \cap B) = P(A|B) \times P(B)}$ → Product rule of prob.

parallely,

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

$\Rightarrow P(A \cap B) = P(A|B)\ P(B) = P(B|A)\ P(A).$

↳ Baye's Theorem.

$$\therefore \boxed{\begin{aligned} P(A|B) &= \frac{P(B|A)\ P(A)}{P(B)} \\ \text{or} \quad P(B|A) &= \frac{P(A|B)\ P(B)}{P(A)} \end{aligned}}$$ → Baye's rule.

we can express $P(A|B)$ in terms of $P(B|A)$.

From the idea of disjoint cover of Ω, we had →

$\{B_1, B_2, \ldots, B_n\}$

$$\left[\begin{aligned} & B_i \cap B_j = \phi \quad \forall\ i \neq j \\ & \bigcup_{i=1}^{n} B_i = \Omega \end{aligned}\right]$$

No B_i is empty

$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots\cdots (A \cap B_n).$ → disjoint union

$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_n)$

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

$$\therefore \boxed{P(A) = \sum_{i=1}^{n} P(A|B_i)\ P(B_i)}$$ → Total Law of probability.

$$P(B_i|A) = \frac{P(A|B_i)\ P(B_i)}{P(A)}$$

$$\therefore \boxed{P(B_j|A) = \frac{P(A|B_j)\ P(B_j)}{\sum_{i=1}^{n} P(A|B_i)\ P(B_i)}}$$

Example → Real life application.

A - A student scores grade A.

B_i - Attendance record of a student (suppose in batches of 10%.)

$P(A|B_j)$ → Historical data used to find likelyhood.

$P(B_i)$ → Prior knowledge of B_i.

The results of baye's theorem are largely dependent on the sample space.

Date:- 14/1/19.

Given 3 events A, B, C.

1) If A & B are independant then,

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

2) If A, B, C are pairwise independant then,

$\boxed{P(C|A \cap B) = 1}$ → Leads to defn of mutual independance.

$$P(A \cap B \cap C) = P(A)\, P(B)\, P(C).$$

* Pairwise Independance Defn:- Let, A, B be 2 events from $(\Omega, \mathcal{F}, \sigma)$. we call A, B pairwise independant of $P(A \cap B) = P(A) \cdot P(B)$.

* Mutually Independant:- If A, B, C are three events from $(\Omega, \mathcal{F}, \sigma)$, we call A, B, C mutually independant if →

$P(A \cap B) = P(A)\, P(B)$ $\quad$ $P(B \cap C) = P(B)\, P(C)$

$P(C \cap A) = P(C)\, P(A)$ $\quad$ $P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$.

we can generalize the mutual independance to the events $A_1, A_2, \dots, A_n$ in an inductive way.

Any sub collection of $A_1, \dots, A_r$ combining atleast 2 events & atmost $(n-1)$ events is mutually independant.

- $P(A_1 \cap A_2 \cap A_3 \dots \cap A_n) = P(A_1) P(A_2) \cdots P(A_n)$.

setting up probability space incase of repeated random experiment (with only 2 possible outcomes in two case).

Experiment → Tossing a coin.

$\Omega = \{0, 1\}$.

$\mathcal{F} = 2^{\Omega} = \{\phi, \Omega, \{0\}, \{1\}\}$.

$P(\{0\}) = 1 - p \qquad P(\{1\}) = p$.

set up a probability space when one repeats R 'n' no. of times with the assumption the repitiens are independant of each other.

For this compound experiment,

$\Omega = \{y_1, y_2, \dots y_n\} \qquad y_i \in \{0, 1\}$.

(sequence of 0, 1).
of length n.

$n(\Omega) = 2^n$.

$\mathcal{F} = \{\ \} \rightarrow$ power set of Ω (2^{Ω}).

★ Probability Assignment :-

Probability of k successes in n repetition at specified location.

so, consider an element $\omega \in \Omega$.

$\omega = y_1, y_2, \dots, y_k, \dots, y_n$.

$y_1, y_2, y_3, \dots, y_k = 1$. $y_{k+1}, y_{k+2}, \dots y_n = 0$.

$$\omega = (\underbrace{1\ 1 \cdots 1}_{k}\ \underbrace{0\ 0 \cdots 0}_{n-k})$$

A_i : success at ith toss for $i = 1, 2, \dots, n$.

$$P(\omega) = P(y_1, \dots y_n) = P(\underbrace{1\ 1 \dots 1}_{k}\ \underbrace{0\ 0 \dots 0}_{n-k.})$$

$$= P(A_1 \cap A_2 \cap A_3 \cdots \cap A_k \cap A_{k+1}^c \cap \cdots \cap A_n^c)$$

$$= P(A_1)\ P(A_2) \cdots P(A_k)\ P(A_{k+1}^c) \cdots P(A_n^c)$$

$$= \underbrace{R \cdot R \cdots R}_{k \text{ times}}\ \underbrace{(1-p)\ (1-p) \cdots (1-p)}_{n-k \text{ times}}$$

$$\therefore P(\omega) = R^k\ (1-p)^{n-k}.$$

★ check → If A, B are mutually independant then, (A, B^c) (A^c, B) & (A^c, B^c) are also mutually independant.

Imp Now if one is interested in computing the probability of the event that exactly k number of success in n coin tosses.

$$\boxed{{}^nC_k\ p^k\ (1-p)^{n-k}}$$

Example → A ~~coin~~ box contains 10 ~~drills~~ balls, 6 are red & 4 are blue.

(A) event 1 → 1st ball picked up is red. &

(B) event 2 → second is blue.

(i) with replacement (ii) without replacement.

As the number of balls increases, the probability of the events without replacement approaches the probability of the events with replacement.

In our eg → For (i) $P(B|A) = 4/10 = P(B)$ (Independant)

ii) $P(B|A) = 4/9$.

$n = 100$. → for (i) $P(B/A) = 40/100 = P(B)$ (Independant)

ii) $P(B/A) = 40/99$

we see $P(B|A)$ tends to $P(B)$ as n increases.

$$P(B) = \frac{6}{10} \times \frac{4}{9} + \frac{4}{10} \times \frac{3}{9}.$$

A & B start behaving like independant events as no. of balls increases.

* Uniform probability spaces:-

exp: Picking up a number randomly from a given set

given set is finite & discrete.

$\Omega = \{1, 2, \dots, n\}$ & any number is equally likely to be picked.

If we pick 2 subsets of Ω which are of equal size, the probability of the 2 events is same.

Imp In case of continuous sample space, a singleton element always has probability 0.

Probability of 2 subsets of Ω with equal size be equal (This is uniform probability principle).

Scanned by CamScanner

Probability of 2 ~~events~~ subsets of Ω with equal size be equal (This is the uniform probability principle).

Here we associate size with cardinality of the set.

Therefore, singleton sets are assigned probability $\frac{1}{n}$.

[These spaces are called as discret uniform probability spaces].

Case 2 → Given set Ω is finite & an interval of R.

In particular take $\Omega = [0,1]$.

* Principle of uniformity:- Pick 2 subsets of Ω with equal size then, they should have same probability.

Take an interval $(a,b) \subset [0,1]$.

size of (a,b) is associated with the length of interval (a,b). which is $(b-a)$.

Probability assignment by oluying principle of uniformity.

This 2 intervals of same length are assigned with probability equal to their length.

The σ- algebra here is the set of subsets of $[0,1]$ which are generated as countable unions & intersections of intervals.

This is continuous uniform probability space.

Notes:-

1. Any singleton set $\{a\} \subset [0,1]$ has probability zero.

2. consider the set of rational numbers in zero to one as $Q \cap [0,1]$.

Then, form (1) & the fact that $Q \cap [0,1]$ is countable

$\hookrightarrow$

$$P(Q \cap [0,1]) = 0.$$

[lecture notes of 15-1-19 are written before 7/1/19].

Khare's class:-

Date:- 15-1-19

7, 8, 14, 15, 21, 22, 28, 29 Jan.

The σ-algebra $\mathcal{F}$, here is the set of ~~statis~~ subsets of $[0,1]$ which are generated as countable unions & intersection of intervals (borel σ-algebra). This is the continuous uniform probability space.

Remarks: 1) Note that any singleton set $\{a\} \subseteq [0,1]$ has prob. zero

2) consider the set of rational no. in $[0,1]$ denoted as $Q \cap [0,1]$.

Then from (ii) and the fact that $Q \cap [0,1]$ in countable.

$P[Q \cap [0,1]) = 0$.

* Problems: -

1. A machine contains four components in parallel with 0.1, 0.2, 0.3, 0.4 as their probability of failures resp. The machine fails if all the components fail simultaneously. Then, note that the failure of machine is independent. Then, what is the probability that the machine once started will not fail.

→ $P = 1 - 0.1 \times 0.2 \times 0.3 \times 0.4 = 1 - 0.0024 = 0.9976.$

soln:

Let, A_i: event that the component i fails.

$A_1 \cap A_2 \cap A_3 \cap A_4 \rightarrow$ machine fails.

Scanned by CamScanner

$= 1 - P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1 \cap A_2 \cap A_3 \cap A_4)^c$.

21

($\because$ failing of machines are independant so,

$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4)$.

2. A rocket engine fails if one key component fails. the probability of failure of those key component is 5%. in order to increase the success probability of the rocket engine & assembly of this key components in parallel is proposed so, that the engine fails if all those key components fails simultaneously. what is the min no. of key components are required in the parallel assembly so that the engine has 99% of success probability.

$\rightarrow$ $0.99 = 1 - P(A_1 \cap A_2 \cdots \cap A_n)$

$= 1 - (P(A_1) \times P(A_2) \times \cdots P(A_n))$.

$0.99 \leq 1 - (0.5)^n$

$(0.5)^n \leq 0.01$

$\therefore$ $n \log_e 0.5 \geq \log_e 0.01$.

$n \geq 6.64$.

$\Rightarrow n \sim 7$.

3. A course on probability is a very famous course & students are allowed to register after the teacher concerns. It is observed that 20% of the times after obtaining the concerns, students fails to register for the course. There are only 100 seats available in the class. If a teacher concerned 102 students what is the prob. that all the students will be accomodated in the class.

→

$$P = 1 - \left[\underbrace{(0.8)^{102}}_{102 \text{ success}} + \underbrace{{}^{102}C_{101} \, (0.8)^{101} \cdot (0.2)}_{101 \text{ success.}} \right].$$

4. In the class of 100 students

MI -	30
IM	40
AG -	30.

70% of MI students are dual degree, 40% of IM students are dual degree, 50% of AG dual degree students

→

5. There are 3 printers P_1, P_2, P_3, P_1 gets 50% of jobs, P_2 gets 30% of jobs, P_3 gets 20% of jobs.

Printer P_1 → has 0.15 as its prob. of failure.

Printer P_2 → has 0.10 —"—.

Printer P_3 → has 0.2 —"—

If a randomly selected printing job is a failure what is the prob. that it was printed by printer 3.

→ P = 1 −

F: event F is failed job → Failed Job.

$P(F|P_1) = 0.15$, $P(F|P_2) = 0.1$, $P(F|P_3) = 0.2$.

$P(P_1) = 0.5$, $P(P_2) = 0.3$, $P(P_3) = 0.2$.

compute, $P(P_3|F)$.

$$P(P_3|F) = \frac{P(F|P_3) \cdot P(P_3)}{\sum_{i=1}^{3} P(F/P_i)\, P(P_i)}.$$

$$= \frac{0.2 \times 0.2}{0.15 \times 0.5 + 0.1 \times 0.3 + 0.2 \times 0.2}$$

$= 0.32$.

6. 2 Fair dice are thrown. what is the prob. that one will obtain two sixes given that

i) one six has already appeared. (1/6)

ii) sum of the two faces is greater than 6. (1/21)

→ A − Appearance of 6 on 1st die.

B − 〃 − on 2nd die.

$$P(B|A) = \frac{P(B \cap A)}{P(B)} = \frac{\frac{1}{6} \times \frac{1}{6}}{\frac{1}{6}} = \frac{1}{6}.$$

---

soln : A = 2 sixes appear.

B = at least one six has appeared.

$\Omega = \{(1,1), (1,2), \ldots (1,6),$

$(6,1), \ldots (6,3), \ldots (6,6)\}.$

Date:- 21- 1-19.

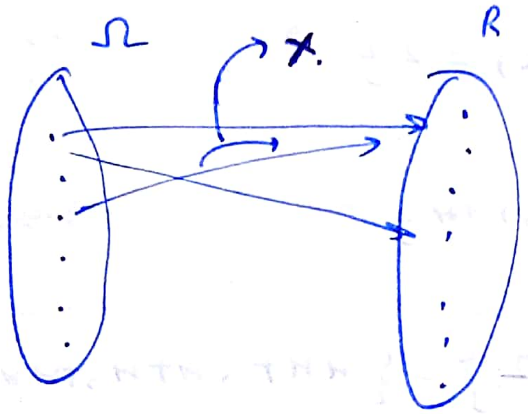
* Random experiment :-

$(\Omega, \mathcal{F}, P)$

* Random variable :-

A random variable X is a function from Ω to R.

if $\{w : X(w) = x\} \in \mathcal{F}$



$\forall\ x \in R$, $\{w : X(w) = x\} = X^{-1}(x)$ [inverse image of X].

$\{w \mid X(w) = x\} \in \mathcal{F}$.

$P(X = x) = P\{w \mid X(w) = x\}$.

e.g.,

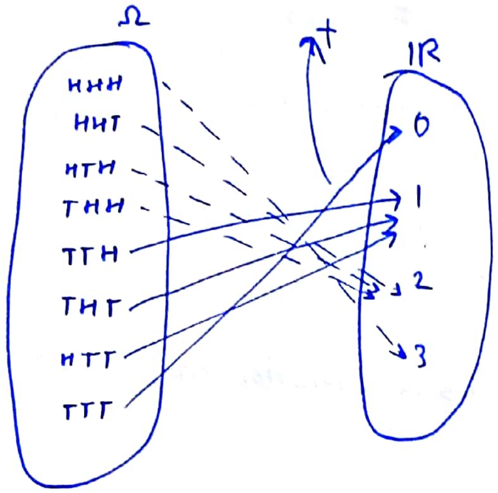
① 3 fair coins tossed independantly.

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

$\mathcal{F} = 2^{\Omega}$.

$P(HHH) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$.

X : no. of heads in 3 tosses.



$x = 2$, $X(2) = \{w \subset \Omega \mid X(w) = 2\}$.

$= \{HHT, HTH, THH\}$.

$P(X=2) = P\{w \subset \Omega \mid X(w) = 2\} = \{HHT, HTH, THH\} = \frac{3}{8}$

$x = 8.7$

$P(X = x) = P(X = 8.87) = 0$.

★ Range of X : $\{x \in R \mid X(x) > 0\}$.

★ A random variable X is called a discrete random variable if range of x is finite or countably infinite

In the previous example:-

$R_x = \{0, 1, 2, 3\}$ & hence x is a discrete random variable.

* Defination:-

A real valued function $f(x) = P(X=x)$ is called as discrete density or probability mass function (Pmf) of the discrete ~~r.v.~~ random variable X.

eg.,

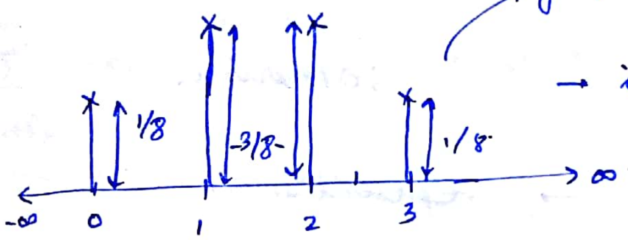
$f(x) = \frac{1}{8}$ $x=0$ $f(x)=0$ for all other value of x except this values

$= \frac{3}{8}$ $x=1$

$= \frac{3}{8}$ $x=2$

$= \frac{1}{8}$ $x=3$

graphical representation of pmf.



→ its nonzero only at discrete points.

representation of function within tabular form.

x	0	1	2	3
f(x)	1/8	3/8	3/8	1/8

→ tabular representation of pmf.

Ex:

① Bernoli random ~~exp.~~ variable. (two types of ~~success~~) outcomes).

expt: Tossing a coin

$\Omega = \{H, T\}$ $P(H) = p$, $P(T) = 1-p$.

$X(\text{Heads}) = 1$, $X(\text{Tails}) = 0$.

$X \sim \text{Bernoli}(p)$ p-is parameter of distribution for

↳ p is the probability of success i.e. $X=1$, $P(X=1)=p$.

$\therefore P(X=0) = 1-p$; $P(X=1) = p$

$$f(x) = \begin{cases} 1-p & ; x = 0 \\ p & ; x = 1. \\ 0 & ; \text{otherwise.} \end{cases} \Bigg\} \text{ pmf.}$$

$R_x \in \{0,1\}$

(2) Binomial random variable:-

expt: n independant bernoli trials are carried out/performed

x : no. of successes in n trials.

$\therefore R_x = \{0, 1, 2, \ldots, n\}$.

$f(x) = P(X=x)$

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} . & ; x = 0, 1, 2, \ldots, n \\ 0 & ; \text{otherwise.} \end{cases}$$

$x \sim$ Bionomial (n, p) $\rightarrow$ representation.

(3) Geometric random variable :-

expt.:- Perform independant bernoli trials (p) until the first success.

$$\Omega = \{ \underset{p}{S}, \underset{(1-p)p}{FS}, \overset{(1-p)^2 p}{FFS}, \underset{(1-p)^3 p}{FFFS}, \ldots \}$$

failure

success.

x : no. of failures ~~get~~ before the first success.

$R_x = \{0, 1, 2, \ldots\}$ $\rightarrow$ discrete random variable.

Scanned by CamScanner

$$f(x) = \begin{cases} p(1-p)^x & ; \quad x = 0, 1, 2, \ldots \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Proof:-

* Note that for any random experiment $P(\Omega) = 1$, Let X be a random variable defined in Ω.

Let R_x be the range of this random variable.

$R_x = \{x_1, x_2, \cdots, x_n, \cdots\}$.

(discrete random variable case).

$$\sum_{x_i \in R_x} F(x_i) = \sum_{x_i \in R_x} P(X = x_i)$$

$$= \sum_{x_i \in R_x} P\{\omega : X(\omega) = x_i\}$$

$$\boxed{\sum_{x_i \in R_x} F(x_i) = 1.}$$ → sum of all pmf is 1

Random expt.: $\Omega, \mathcal{F}, P \to X \to$ pmf.

* Properties of pmf:-

1. $f(x) \geq 0 \quad ; \forall x \in R$
2. $\{x : f(x) > 0\}$ is a finite or countable set.
3. $\sum_{x_i \in R_x} f(x_i) = 1$

If a function f satisfy all these three properties then, the function is pmf. or discrete density.

④ Uniform density: or uniform pmf :-

$$f(x) = \begin{cases} \frac{1}{s} & ; x \in \{x_1, x_2, \dots x_s\}. \\ 0 & \text{otherwise}. \end{cases}$$

⑤ Geometric pmf:-

$$f(x) = \begin{cases} p(1-p)^x & ; x = 0, 1, 2, \dots \\ 0 & ; \text{otherwise}. \end{cases}$$

⑥ Negative binomial distribution : (α & p) are the parameters for this distribution.

$$f(x) = \begin{cases} p^\alpha \binom{\alpha+x-1}{x} (1-p)^x & x = 0, 1, 2, \dots \\ 0 & ; \text{otherwise} \end{cases}$$

(expt. performing independent bernoli (p) trials unit α no. of success are observed).

for $\alpha = 2$. x :- no. of failures before α success $\Rightarrow$ length of string $= \alpha + x$.

$\{HH, HTH, THH, \dots\}$.

⑦ Poisson density (λ) :- $\lambda > 0$

$$f(x) = \begin{cases} \frac{e^{-\lambda} \cdot \lambda^x}{\lfloor x} & ; x = 0, 1, 2, \dots \\ 0 & ; \text{otherwise} \end{cases}$$

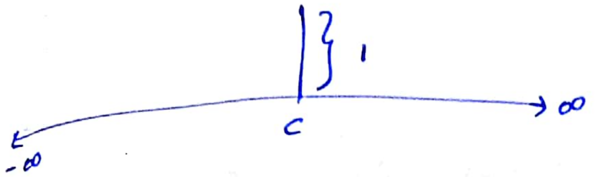
Scanned by CamScanner

* special examples of random variables:-

① Ex: constant random variable

$X(\omega) = c \quad \forall\ \omega \in \Omega.$

$P(X = c) = 1, \quad P(X \neq c) = 0.$



② Ex: Indicator random variable:-

let, $A \subset \mathcal{F}$.

$X_A(\omega) = 1 \qquad \forall\ \omega \in A.$

$\quad = 0 \qquad \forall\ \omega \in A^c.$

Bernoli random variable is basically a Indicator random variable.

③ Ex:- Hypergeometric distribution:- doubt

Population of r objects.

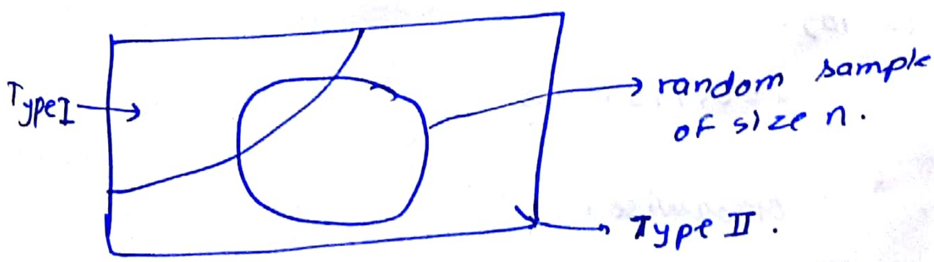
Type I objects r_1

Type II objects $r - r_1 = r_2$.

Let a sample of size n is chosen from this population $(n \leq r)$.

$n < r_1$

X: no. of objects of type I in the random sample.



$$P(X=x) = \frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}} \quad ; x = 0, 1, 2, \dots, n \}$$

$$= 0 \quad ; \text{otherwise.}$$

★ **Computation with pmf:-**

$(\Omega, \mathcal{F}, P)$.

X : discrete r.v (random variable) with pmf $f(x)$.

Event: $x \leq t$ for some real no. t.

$= \{ \omega : X(\omega) \leq t \}$ → set of ω such that $X(\omega) \leq t$.

$$= \bigcup_{i=-\infty}^{t} \{ \omega : X(\omega) = i \}.$$

$$P(X \leq t) = P\left(\bigcup_{i=-\infty}^{t} \{ \omega : X(\omega) = i \} \right) = \sum_{i=-\infty}^{t} P(X = i)$$

★ **Cumulative distribution function:- [CDF]**

Discrete random variable X with pmf $f(x)$.

for every real number $t \in \mathbb{R}$, define

$$\left.\begin{aligned} F(t) &= P(X \leq t) \\ &= \sum_{x \leq t} f(x) \end{aligned}\right\} \text{cumulative distribution function.}$$

eg.,

$X \sim \text{uniform}(s = 10)$.

$$f(x) = \frac{1}{10} \quad x = 0, 1, 2, \dots, 9$$

$$= 0 \quad \text{otherwise.}$$

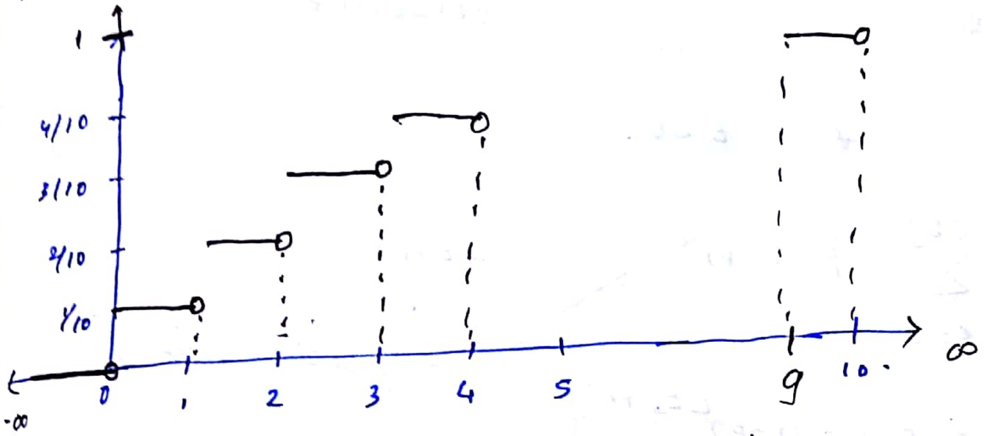
clearly for $t < 0$, $F(t) = 0$.

$F(0) = f(0) = \frac{1}{10}$.

for any $0 \leq t < 1$, $F(t) = \frac{1}{10}$.

for $F(1) = f(0) + f(1) = \frac{1}{10} + \frac{1}{10} = \frac{2}{10}$.

for any $1 \leq t < 2$, $F(t) = \frac{2}{10}$.



CDF: Cumulative distribution Function.

Let X be a discrete random variable with pmf $f(x)$.

For any $t \in R$ we define.

$$F(t) = \sum_{x \leq t} f(x)$$

clearly F is a function from R to R.

1) F is a non-decreasing Function.

For $t_1 < t_2 \Rightarrow F(t_1) \leq F(t_2)$.

2) $F(-\infty) = 0$, $F(\infty) = 1$.

$\Rightarrow \lim_{t \to -\infty} F(t) = 0$, $\lim_{t \to \infty} F(t) = 1$

3) The graph of F is like a staircase with jumps occuring exactly at the points which are in the range of X.

4) Right continuity of F

$$\lim_{x \to a^+} F(x) = F(a) \quad \forall a \in \mathbb{R}.$$

Ex:-

$X \sim$ geometric (p)

$$f(x) = \begin{cases} p(1-p)^x & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

$F(t) = 0 \quad \forall \quad t < 0.$

$$F(t) = \sum_{x=0}^{[t]} p(1-p)^x \qquad t \geq 0$$

geometric series sum.

$$= p\left[\frac{1-(1-p)^{[t]+1}}{1-(1-p)}\right]$$

$$F(t) = [1-(1-p)^{[t]+1}]$$

For this geometric random variable

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

$= 1 - P(X \leq x-1)$

$= 1 - F(x-1).$

$= 1 - [1-(1-p)^{x-1+1}]$

$P(X \geq x) = (1-p)^x.$

events

$X < x$ & $X \geq x.$

$X < x \cup X \geq x = \Omega.$

disjoint.

$\{w : X(w) < x\}$

$\{w : X(w) \geq x\}$

Prove:

$P(X > n+m \mid X > n) = P(X > m)$

↳ memoryless property.

Date:- 22-1-19.

Prove → $P(x > n+m \mid x > n) = P(x > m)$.

↳ Memory less property of

$$\frac{P(x > n+m \cap x > n)}{P(x > n)}$$

$$= \frac{P(x > n+m)}{P(x > n)}$$

$x > n$ → n. no. of tails have already occured.

$x > n+m$ → $n+m$ no. of failures occur

The results tell us that even after knowing that n failures have occured we cant commet in the next event.

prove of getting 15 failures given 10 failures have occured.

The memory less property says that this if is forgotten, & it is equivalent to starting a fresh.

---

x is a random variable $(\Omega, \mathcal{F}, P)$.

Probability of the event $(X \le x)$

$F(x) = P(X \le x)$ $\forall x \in R$

This function F is called as cumulative distribution function $F(x)$ of the random variable x. $(F_x(x))$

* **Properties of CDF:-**

1. $0 \le F_x(x) \le 1 \quad \forall x \in R.$

2. $\lim_{x \to \infty} F_x(x) = 1.$

3. $\lim_{x \to -\infty} F_x(x) = 0.$

4. Function $F_x(x)$ is non decreasing.

   (for $x_1 \le x_2$, $f(x_1) \le f(x_2)$ $\longrightarrow$ Non-dec.)

5. Right continuity $\rightarrow$

   $\lim_{x \to a^+} F_x(x) = F_x(a) \quad \forall a \in R.$

   For every a in R & $\delta > 0.$

   $\lim_{\delta \to 0} (F_x(a+\delta) - F_x(a)) = 0$

* **Lebesgue Decomposition theorem:-**

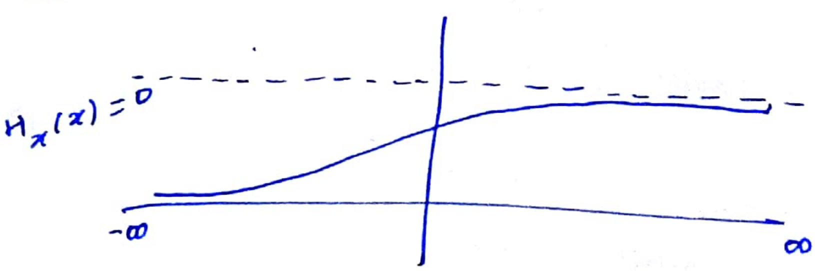
If a function $F_x(x)$ satiesfies the properties stated above, then $F_x(x)$ can be represented as sum of two functions say $G_x(x)$ & $H_x(x)$ $\rightarrow$

$F_x(x) = G_x(x) + H_x(x).$

where $G_x(x)$ is continuous & $H_x(x)$ is a right continuous step function which jumps coinciding with those of $F_x(x)$

$H_x(-\infty) = 0.$

ex.



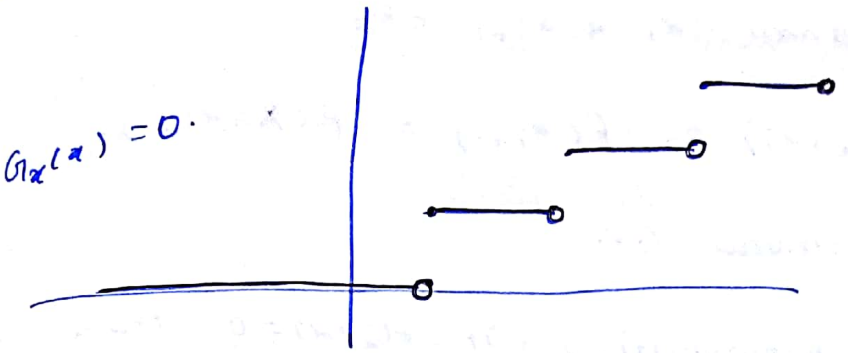
Continuous random variable

$\Rightarrow$ CDF.

ex:-

Discrete r.v.

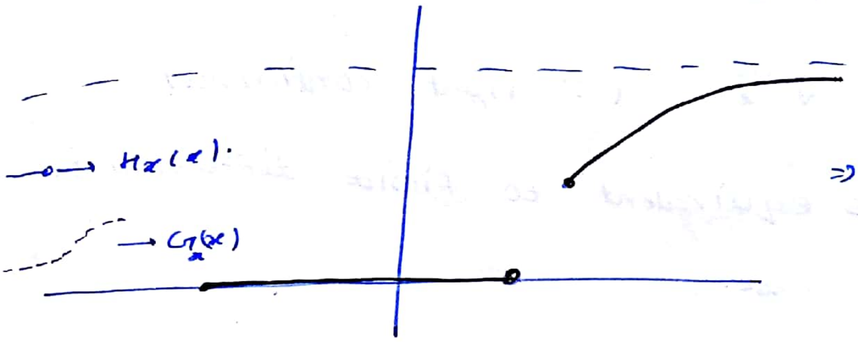
$\Rightarrow$ CDF.



ex:-

Mixed r.v.

$\Rightarrow$ CDF



In lebesgue decomposition $G_x(x) = 0$ then, the r.v. (x) is called as discrete r.v.

If $H_x(x) = 0$, the r.v. (x) is called as continuous r.v.

The situation where neither of these are identically equal to zero; we take r.v. (x) as a mixed random variable.

The example is a normal cdf. (conti.).

• **Defination:-**

If x is a discrete r.v.

$$P(x_i) = f(x_i) = P(x = x_i) > 0 \quad \forall \; x_i \in R_x.$$

1. $\sum f(x_i) = 1.$

2. $F_x(x_i) = P(X \leq x_i) = \sum_{x \leq x_i} f(x_i).$

Let x_i & x_{i+1} $\in R_x$ such that $x_i < x_{i+1}$ & there is no point between x_i & x_{i+1} $\in R_x$

$$F_x(x_{i+1}) - F_x(x_i) = f(x_{i+1}) = P(X = x_{i+1}).$$

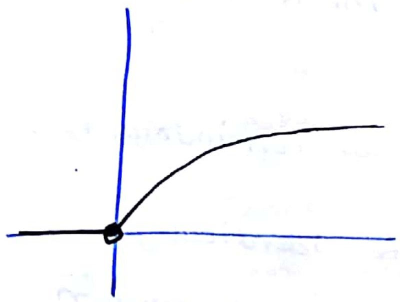
* Defination:- Continuous r.v.

In Lebsegue decomposition, if $H_x(x) = 0$ then, x is a continuous r.v.

$$P(X=x) = \lim_{\delta \to 0^+} [F_x(x+\delta) - F_x(x)]$$

$$= 0 \quad \forall x \quad (\because \text{ right continuous}).$$

The derivative is equivalent to finite difference of discrete variable case.



The derivative is piece wise continuous & has only finitely many jumps

define $f_x(x) = \frac{d}{dx} F_x(x)$.

↑

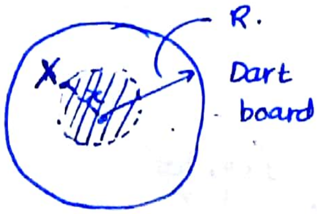
always exists except for a few points.

in the above example $f_x(0)$ does not exist

$f_x(x)$ = probability density function.

Date:- 28-1-19.

example:-



expt:- throw a dart on this circular board.

$(\Omega, \mathcal{F}, P) \rightarrow$ uniform probability space.

Random variable 'x':

x: dist. of the dart from the centre of the board. (bulls eye) 0.

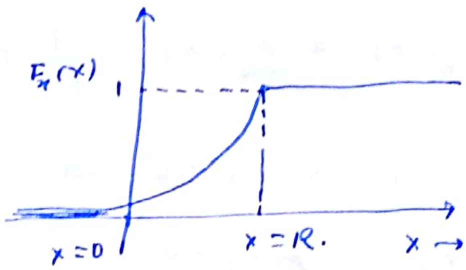
let, $F_x(x)$ denote the cdf of X.

$F_x(x)$ = prob. $(X \leq x)$.

$= \frac{\pi x^2}{\pi R^2}$ (due to uniformity principle).

$= \boxed{x^2/R^2}$

$$F_x(x) = \begin{cases} 0 & , x < 0 \\ x^2/R^2 & , 0 \leq x \leq R \\ 1 & , x > R \end{cases}$$



Probability density function:-

$$f_x(x) = \frac{d}{dx} F_x(x) = \begin{cases} 0 & , x < 0 \\ \frac{2x}{R^2} & , 0 \leq x \leq R \\ 0 & , x > R. \end{cases}$$

$\because f_x(x) = \frac{d}{dx} F_x(x) \Rightarrow F_x(x) = \int_{-\infty}^{x} f_x(x)\, dx.$

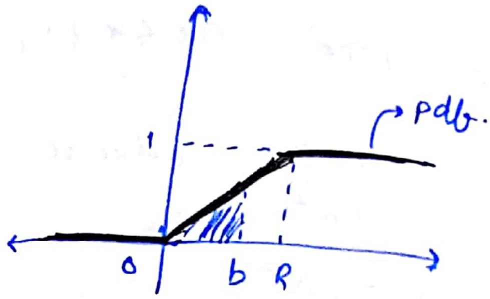
$$f_x(x) = \begin{cases} 2x/R^2 & , 0 \leq x \leq R \\ 0 & , \text{otherwise.} \end{cases}$$

- For continuous pdf does not give prob.

$$P(a \leq x \leq b) = F_X(b) - F_X(a).$$

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

$$= \int_{a}^{b} f_X(x)\,dx.$$



Let x be a continuous random variable with cdf $F_X(x)$ & $f_x(x) = \frac{d}{dx} F_x(x)$ as its pdf : then,

range of X is the set $R_x \in R$ such that

$\forall\ x \in P_X$, $f(x) > 0$

★ **Defination:-**

A density function or pdf $f(x)$ is a non-negative function s.t.

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

obviously,

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

satisfy all properties of CDF.

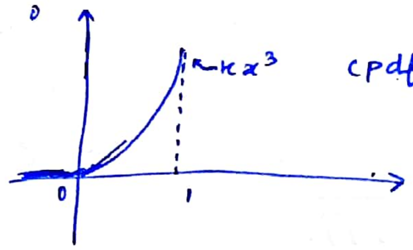
$$P(a \leq x \leq b) = \int_{a}^{b} f_X(x)\,dx = F_X(b) - F_X(a)$$

ex ;

1) For what value of K, $f(x) = \begin{cases} Kx^3 & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$

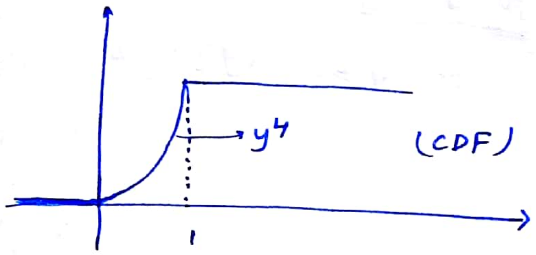
* If x is a random variable, (any func^n) $f(x)$ is also a random variable.

$\rightarrow \int_0^1 kx^3\,dx = 1 \implies k = 4.$



(pdf) f_x.

For $y \in (0,1)$ $\quad F_x(y) = \int_0^y 4x^3\,dx = y^4.$



(CDF) F_x

eg:-

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

* change of variable formula:-

Let X be a continuous r.v. with pdf $f_X(x)$.
Find the density of the random variable $Y = X^2$.

Let, $G_Y(y)$ be the CDF of Y.
let, $F_X(x)$ be the CDF of X.

$G_Y(y) = P(Y \le y)$ $\longrightarrow$ for any real no. $y \in \mathbb{R}$.

$\quad\quad = P(X^2 \le y).$

$G_Y(y) = P(-\sqrt{y} \le X \le \sqrt{y}).$

$\therefore G_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$

Scanned by CamScanner

In order to find density of Y say $g(y)$, we take derivative of $G_Y(y)$.

$$g(y) = \frac{d}{dy} G_Y(y).$$

$$= \frac{d}{dy}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right]$$

$$= f_x(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} + f_x(-\sqrt{y}) \frac{1}{2\sqrt{y}}.$$

$$\boxed{\therefore g(y) = \frac{f_x(\sqrt{y}) + f_x(-\sqrt{y})}{2\sqrt{y}}.}$$ for $y > 0$.

for $y \le 0$, $g(y) = 0$.

ex:

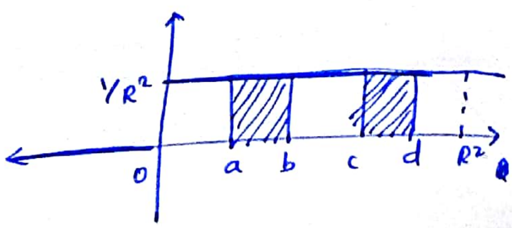
$$f_X(x) = \begin{cases} 2x/R^2 & , 0 \le x \le R \\ 0 & , \text{otherwise}. \end{cases}$$

Find the density of $Y = X^2$, Let $g(y)$ be the density of Y

$$g(y) = \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_x(-\sqrt{y})\right].$$

$$= \frac{1}{2\sqrt{y}}\left[\frac{2\sqrt{y}}{R^2} + 0\right] \quad , 0 \le y \le R^2.$$

$$g(y) = \left.\begin{cases} 1/R^2 & , 0 \le y \le R^2 \\ 0 & , \text{otherwise} \end{cases}\right\}$$ uniform continuous density



any two intervals of same length $(b-a)=(d-c)$ then, the probability is same.

* Uniform probability principle:-

Let, x be a continuous random variable with the pdf defined as

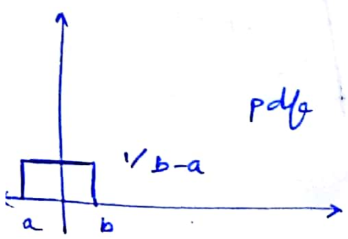
$$f_X(x) = \begin{cases} \frac{1}{b-a} & , a \le x \le b. \\ 0 & , \text{otherwise}. \end{cases}$$

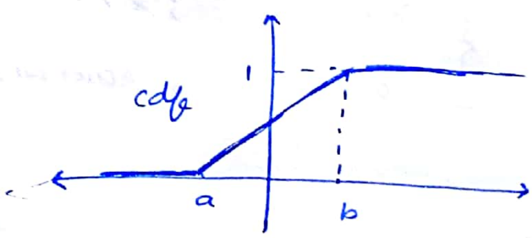
for any two $a, b \in R$ such that $a < b$.

Then, X is said to follow uniform density with parameter a & b. $X \sim U(a,b)$.

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \frac{x-a}{b-a} \; ; \; a < x < b.$$

$$\therefore F_X(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , a \le x \le b \\ 1 & , x > b. \end{cases}$$





ex:-

Let $X \sim U(0,1)$ (uniform r.v.).

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

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

consider the transformation:-

$$Y = -\frac{1}{\lambda} \log(1-X) \quad (\text{for } \lambda > 0)$$

Let, $G_Y(y)$ be the CDF of Y & $g(y)$ be the pdf of Y.

$$G_Y(y) = P(Y \le y)$$

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

$$= P[\log(1-X) \geq -\lambda y].$$

$$= P[(1-X) \geq e^{-\lambda y}].$$

$$= P[X \le 1-e^{-\lambda y}]$$

$$\therefore G_Y(y) = 1-e^{-\lambda y}.$$

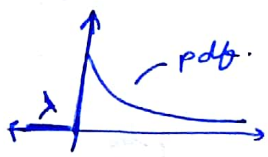
$$g(y) = \begin{cases} \lambda e^{-\lambda y} & y > 0 \\ 0 & y < 0. \end{cases}$$

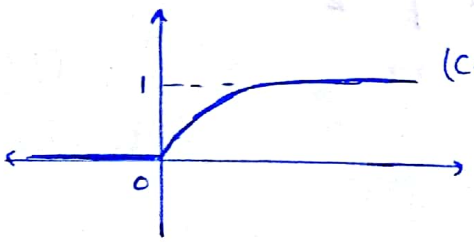
This is an exponential density with parameter $(\lambda > 0)$

Exercise:

Let, $X \sim \exp(\lambda)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & ; x > 0 \\ 0 & ; \text{otherwise} \end{cases}$$





cdf:

$$F_X(x) = \begin{cases} 0 & ; x < 0 \\ 1-e^{-\lambda x} & ; x \geq 0. \end{cases}$$

For any real no. $x, y > 0$.

$$P(X > t) = F_X(\infty^-) - F_X(t)$$

$$= 1-(1-e^{-\lambda t}) = e^{-\lambda t}.$$

$P(X > x) = e^{-\lambda x}$ $\qquad [P(X > t) = e^{-\lambda t}]$.

$P(X > y) = e^{-\lambda y}$

$P(X > (x+y)) = e^{-\lambda(x+y)} = P(X > x) \cdot P(X > y)$.

$\therefore \dfrac{P(X > (x+y))}{P(X > y)} = P(X > x)$.

$$\boxed{P(X > (x+y) \mid X > y) = P(X > x)}$$

Memoryless property of exponential density.

*Theorem:-

Let ϕ be a differentiable function which is strictly increasing or strictly decreasing on an interval I. Let, $\phi(I)$ denote the range of ϕ & ϕ^{-1} be the inverse of ϕ on I. Let X be a continuous r.v. having density $f_X(x)$ s.t. $f_X(x) \neq 0$ on I.

Then, $Y = \phi(X)$ whose density is given by $g(y)$ &

$$g(y) = f(\phi^{-1}(y)) \left| \frac{d}{dy}(\phi^{-1}(y)) \right| \quad [y \in \phi(I)]$$

mod.

*Proof:-

Let $G(y)$ be the CDF of y.

$G(y) = P(Y \le y)$.

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

$= P(X \le \phi^{-1}(y))$ $\quad$ } f is increasing.

$G(y) = F_X(\phi^{-1}(y))$.

$\Rightarrow g(y) = F_X'(\phi^{-1}(y)) \dfrac{d}{dy}(\phi^{-1}(y))$

$$\therefore \boxed{g(y) = f_X(\phi^{-1}(y)) \frac{d}{dy}(\phi^{-1}(y))} \quad (1)$$

ϕ is a decreasing function.

$G(y) = P(Y \le y)$

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

$= P(X \geq \phi^{-1}(y))$ } ϕ is decreasing.

$G(y) = 1 - F_X(\phi^{-1}(y))$

$$\boxed{g(y) = -f_X(\phi^{-1}(y)) \frac{d}{dy}(\phi^{-1}(y))} \quad — ②$$

combining ① & ② results,

$$\Rightarrow \boxed{g(y) = f_X(\phi^{-1}(y)) \left| \frac{d}{dy}(\phi^{-1}(y)) \right|}$$

ex:

① Let X be a random variable with density f. Let, $a, b \in R$ & $b \neq 0$

Then, define $Y = a + bX$ then what is the density of Y.

$\rightarrow$ $g(y) = f(\phi^{-1}(y)) \left| \frac{d}{dy}(\phi^{-1}(y)) \right|$.

$\phi(x) = a + bx = y$.

$x = \frac{y-a}{b} = \phi^{-1}(y)$.

$\therefore g(y) = f\left(\frac{y-a}{b}\right) \left|\frac{1}{b}\right|$.

② $f_X(x) = \begin{cases} 2x/R^2 & ; 0 < x < R \\ 0 & ; \text{otherwise.} \end{cases}$

$\rightarrow$ $Y = \frac{x}{R}$ $\left[a = 0, b = \frac{1}{R}\right]$ (linear transformation).

$x = YR$

$\Rightarrow 0 < YR < R$.

$0 < Y < 1$ $\longrightarrow$ range of Y.

$$\cancel{g(y) = \begin{cases} 2y & ; \\ 0 & \end{cases}}$$

$$g(y) = \frac{1}{|b|} \quad f\left(\frac{y-a}{b}\right)$$

$$= \frac{1}{1/R} \cdot \frac{2y/(1/R)}{R^2}$$

$$\therefore g(y) = \begin{cases} 2y & ; \quad 0 \le y \le 1 \\ 0 & ; \quad \text{otherwise.} \end{cases}$$

* symmetric densities:-

f is symmetric density if $f(x) = f(-x) \quad \forall x \in R.$

$U(-a, a)$ is symmetric.

A random variable is symmetric if its pdf $f_X(x)$ is symmetric density.

Ex:-

If X is a symmetric random variable with CDF $F_X(x)$ Then, $F_X(0) = \frac{1}{2}$.

~~replacing~~ putting ~~y as~~ -y

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

$$= \int_{x}^{\infty} f(y)\,dy \quad \Big) \because f(y) = f(-y).$$

$$= 1 - F_X(x)$$

$$F_X(-x) = 1 - F_X(x). \quad \forall x \in R.$$

ex:- $-\infty < x < \infty$.

$$g(x) = \frac{1}{1+x^2}$$

Is $g(x)$ a pdf ??

$$\rightarrow \int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \tan^{-1} x \Big|_{-\infty}^{\infty} = \pi.$$

$-\infty < x < \infty$.

$$\boxed{\therefore g(x) = \frac{1}{\pi(1+x^2)}}$$

is a density called cauchy density

since $g(x) = g(-x)$; cauchy density is symmetric.

Ex:

$g(x) = e^{-x^2/2}$ $\quad -\infty < x < \infty$.

$$c = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx.$$

$$c^2 = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \cdot \int_{-\infty}^{\infty} e^{-x^2/2}\,dx.$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,dy\,dx.$$

by polar co-ordinates substitution,

$x = r\cos\theta$, $y = r\sin\theta$.

$$\therefore c^2 = \int_{-\infty}^{\infty}\int_{0}^{2\pi} r e^{-r^2/2}\,d\theta\,dr \qquad \Rightarrow c^2 = 2\pi \quad \Rightarrow c = \sqrt{2\pi}$$

$$\cancel{= -2\pi \int_{-\infty}^{\infty} -r e^{-r^2/2}\,dr} = \cancel{-2\pi \left[e^{-r^2/2}\right]_{-\infty}^{\infty}}$$

$$f^2 = 2\pi \int_{-\infty}^{\infty} r e^{-r^2/2} dr. =$$

$$\therefore \int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

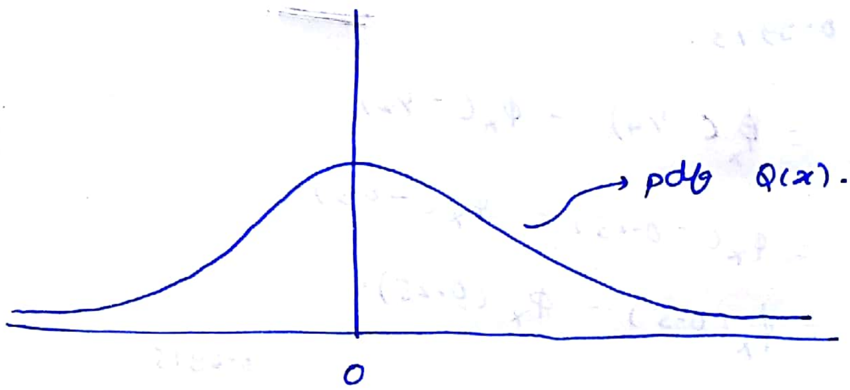
Define

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad -\infty < x < \infty.$$

clearly this $\phi(x)$ is a pdf. This density is called as standard normal density.

The random variable corresp. to this density X, is called as standard normal r.v.

So, the $\boxed{\text{standard normal density}}$ $\phi(x)$ is symmetric.



$R_X = R$ for $\phi(x) > 0$, $\forall x \in R$.

$P(a \leq X \leq b) = F_X(b) - F_X(a)$.

Let, Φ_X denote the CDF of standard normal r.v.

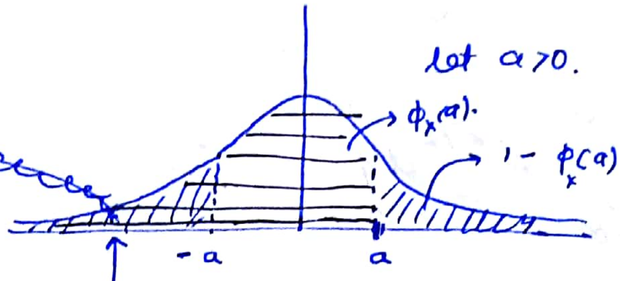
$$\therefore P(a \leq X \leq b) = \Phi_X(b) - \Phi_X(a)$$

$$= \int_a^b \phi(x) dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

It seems that it is enough to know

$\phi_X(a) \quad \forall a \in \mathbb{R}.$

$$\phi_X(x) = \int_{-\infty}^{x} \phi(x)\, dx$$



$\because$ for symmetric densities,

$F_X(-x) = 1 - F_X(x)$

$\therefore \phi_X(-x) = 1 - \phi_X(x)$

evaluate $\phi_X(-a)$

1. $P(-3 < X < 3) = \phi_X(3) - \phi_X(-3).$

$= 2\phi_X(3) - 1$

$= 2 \times 0.9986 - 1.$ $\longrightarrow$ From standard normal density table.

$= 0.9973.$

2. $P(-\frac{1}{2} < x < -\frac{1}{4}) = \phi_X(-1/4) - \phi_X(-1/2)$

$= \phi_X(-0.25) - \phi_X(-0.5)$

$= \phi_X(0.5) - \phi_X(0.25).$

~~= 0.7224~~

$= 0.6915 - 0.5987$

$$\begin{array}{r} 0.6915 \\ 0.5987 \\ \hline 0.0928 \end{array}$$

$= 0.0928.$

- Let X ~~follow~~ be a standard normal r.v.

Define $Y = \mu + \sigma X$ where $\sigma > 0$.

Find density of Y.

let, $g(y)$ be the density

$$\therefore g(y) = \frac{1}{\sigma}\, \phi\left(\frac{y-\mu}{\sigma}\right).$$

$$\therefore g(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \qquad -\infty < y < \infty.$$

This Y is said to follow normal density with two parameters μ & σ^2.

$Y \sim N(\mu, \sigma^2)$.

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \qquad -\infty < y < \infty.$$

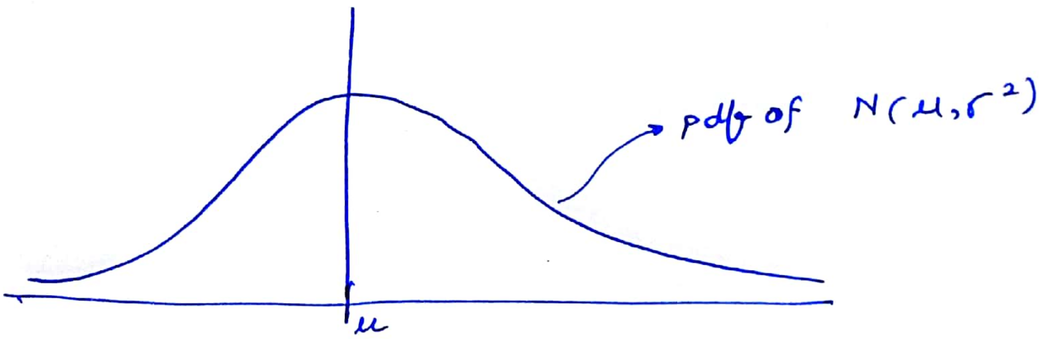
* Probability computation:-

$$P(a < Y < b) = P(a < \mu + \sigma X < b)$$

$$= P(a - \mu < \sigma X < b - \mu).$$

$$= P\left(\frac{a-\mu}{\sigma} < X < \frac{b-\mu}{\sigma}\right).$$

$$= \phi_X\left(\frac{b-\mu}{\sigma}\right) - \phi_X\left(\frac{a-\mu}{\sigma}\right).$$



$Y \sim N(\mu, \sigma^2)$

$$P(\mu - 3\sigma < Y < \mu + 3\sigma).$$

$$= P(\mu - 3\sigma - \mu < Y - \mu < \mu + 3\sigma - \mu)$$

$$= P\left(\frac{-3\sigma}{\sigma} < \frac{Y-\mu}{\sigma} < \frac{3\sigma}{\sigma}\right)$$

$$= P(-3 < X < 3)$$

$$= \phi_X(3) - \phi_X(-3) = 2\phi_X(3) - 1$$

$$= 0.9970.$$

→ Gamma densities :-

Let $X \sim N(0, \sigma^2)$

$Y = X^2$

Find density of Y.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \quad -\infty < x < \infty$$

Let $g(y)$ = density of y.

$$g(y) = \frac{1}{2\sqrt{y}} \left(f(\sqrt{y}) + f(-\sqrt{y}) \right) \quad y > 0$$

Gamma density

$$g(y) = \frac{1}{\sigma\sqrt{2\pi y}} \left[e^{-y/2\sigma^2} \right] \quad y > 0.$$

Gamma func^n's :-

$$\int_0^\infty x^{\alpha-1} e^{-\lambda x} \, dx = \frac{\Gamma(\alpha)}{\lambda^\alpha}$$

Gamma pdf :-

$$\Gamma(x;\alpha,\lambda) = \begin{cases} \dfrac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x>0 \\ 0, & x \le 0. \end{cases}$$

Parameters of this density are (α, λ).

Take $\alpha = 1/2$, $\lambda = 1/2\sigma^2$

EXPECTATION OF A RANDOM VARIABLE :-

r.v.s.

discrete | ~~random~~ continuous

- R_x countably finite / infinite.
- p.m.f

- p.d.f.

Def^n :- Let X be a discrete random variable with R_x as its range.

$$\boxed{E(X) = \sum_{x \in R_X} x P(X=x)}$$ (provided the sum is convergent)

Ex: Let X be a ~~random~~ discrete uniform r.v. on $\{x_1, x_2, \dots, x_n\} = R_X$

$$P(X=x_i) = 1/n \rightarrow \text{for } x_i \in R_X$$
$$= 0 \quad \text{otherwise}$$

$$E(X) = \sum_{x_i \in R_X} x_i P(X=x_i) = \sum x_i p(x_i) = \frac{\sum x_i}{n} = \text{A·Mean.}$$

In case of discrete uniform r.v. $E(X)$ is the A.M. of R_X

In general, we can understand $E(X)$ as the weighted average

Ex: $\cancel{X} \sim$ Bernouli(p)

$X \sim$ Berouli (p)

X	0	1
$p(X)$	$1-p$	p

← pmf of X in tabular form

$$E(X) = 0 \cdot (1-p) + 1 \cdot p = p$$

Ex: $X \sim$ Binomial (n, p)

$$p(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & , x = 0, 1, \ldots, n \\ 0 & , \text{otherwise.} \end{cases}$$

$$E(X) = \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j}$$

notice: $j\binom{n}{j} = \frac{j \cdot n!}{j!(n-j)!} = n\binom{n-1}{j-1}$

$$E(X) = np \sum_{j=1}^{n} \binom{n-1}{j-1} p^{j-1} (1-p)^{n-j+1}$$

$$= np\,(p + (1-p))^{n-1} \quad \text{– (From Binomial expansion)}$$

$$= np.$$

$\boxed{E(X) = np}$ Prove.

Ex: $X \sim \text{Poisson}(\lambda)$

$$p(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & x = 0,1,2,\ldots \\ 0 & \text{o.w.} \end{cases}$$

$$E(X) = \sum_{j=0}^{\infty} j \frac{e^{-\lambda}\lambda^{j}}{j!}$$

$$\boxed{E(X) = \lambda}$$

Ex:- $X \sim \text{Geometric}(p)$ → (no. of failures before one success)

$$p(X=x) = \begin{cases} (1-p)^{x} p & , (x=0,1,2,\ldots) \\ 0 & , \text{otherwise.} \end{cases}$$

$$E(X) = \sum_{j=0}^{\infty} j\, p(1-p)^{j}$$

$$= p(1-p) \sum_{j=0}^{\infty} j (1-p)^{j-1}$$

$$= -p(1-p) \sum_{j=0}^{\infty} \frac{d}{dp} (1-p)^{j-1}$$

Interchanging sum and d/dp as series absolutely summable.

$$= -p(1-p) \frac{d}{dp} \sum_{j=1}^{\infty} (1-p)^{j-1}$$

$$= -p(1-p) \frac{d}{dp}\left(\frac{1}{p}\right)$$

$$\boxed{E(X) = \frac{(1-p)}{p}}$$

Example - Let X be a discrete r.v. with the p.m.f

$$f(x) = \begin{cases} \frac{1}{x(x+1)} , & x = 1, 2, \ldots \\ 0 , & \text{otherwise} \end{cases}$$

$f(x) \geq 0$ for all $x \in R_X$

and $\displaystyle\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} \frac{1}{x(x+1)} = \sum_{x=1}^{\infty} \left[\frac{1}{x} - \frac{1}{x+1}\right]$

$$= 1 - \cancel{\frac{1}{2}} + \cancel{\frac{1}{2}} - \cancel{\frac{1}{3}} + \cdots \cancel{\frac{1}{n-1}} - \frac{1}{n}$$

$$= 1 - \frac{1}{n} \quad (\text{for } n \text{ finite})$$

$$\text{for } \lim_{n \to \infty} \left(1 - \frac{1}{n}\right) = 1$$

$$\therefore \boxed{\sum_{x=1}^{\infty} f(x) = 1}$$

Thus $f(x)$ is a pmf.

$$E(X) = \sum_{x=1}^{\infty} x \cdot f(x) = \sum_{x=1}^{\infty} \frac{1}{x+1}$$

(unbounded)

E(X) does not exist

(This will happen mostly where ∞ sums come into picture)

~~PROPERTIES OF EXPECTATION~~ :-

Transformation - Expectation of func^n of discrete r.v.:

(1) Let $Z = \phi(x)$ be a func^n of discrete r.v. 'x'

$$E(Z) = E(\phi(x)) = \sum_x \phi(x)\, P(X=x)$$

provided expectation exists

Ex: X a discrete r.v. with pmf

X	-2	-1	0	1	2
p(x)	1/5	1/5	1/5	1/5	1/5

$E(X) = 0$, $E(X^2) = 2$

→ Properties of expectation :- (for Discrete)

Let X and Y be discrete r.v.s. with finite expectations.

(1) For an $\alpha \in \mathbb{R}$ such that $P(X=\alpha) = 1$.

$$\boxed{E(X) = \alpha}$$

(2) For some constant 'c' belonging to $\mathbb{R}$

$$E(cX) = c\,E(X)$$

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

(4) If $P(X \geq Y) = 1$ then $E(X) \geq E(Y)$
moreover, $E(X) = E(Y)$ if and only if $P(X=Y) = 1$.

(5) $|E(X)| \leq E(|X|)$

* **Exercise :** Hypergeometric distribution.

Population (r)

Type I $\quad$ Type II

$r_1 \quad r - r_1$

A sample of size 'n' is drawn from the population.

S_n = no. of objects of Type I in this sample.

Define ; X_i : ith indicator r.v. indicating whether the ith object in the sample is of Type I.

$$S_n = X_1 + X_2 + \dots + X_n$$

$$E(X_i) = r_1/r$$

$$E(S_n) = n\,E(X_i)$$

$$\boxed{E(S_n) = \frac{n r_1}{r}}$$

Moments : $E(X^r) = \sum_x x^r p(X=x)$ for $x = 1, 2, \dots$

These are called as raw moments of X.

$r = 1 \Rightarrow E(X)$ also called as mean of X.

$$E((X-a)^r) = \sum_x (x-a)^r p(X=x)$$

for some fixed $a \in \mathbb{R}$. These are

[*Calculate Variance of the (distributions (r.v.))]

called Central Moments of X.

Take $a=0 \Rightarrow$ Central Moments are same as Raw moments.

In particular take $a = \mu = E(X)$

For $r=2$

$E(X-\mu)^2 = E(X-E(X))^2$ = variance of X

The positive square root of variance of X is called as standard deviation of X.

$E(X-\mu)^2 = \sigma^2 \leftarrow$ notation.

Interpretations of $\sigma^2 = var(X)$:-

Measuring squared variability of X around 'a'.

$= E(X-a)^2$

Interested in minimising $E(X-a)^2$ w.r.t. 'a'.

(Approximating a r.v. by a constant 'a' and the error in the approximation is $E(X-a)^2$).

$$E(X-a)^2 = E(X^2 - 2aX + a^2)$$
$$= E(X^2) - 2aE(X) + a^2$$

differentiating w.r.t. 'a'. $=0$.

$$-2E(X) + 2a = 0$$
$$\Rightarrow \boxed{a = E(X)} = \mu \text{ (for minimising)}$$

Scanned by CamScanner

Another interpretation of variance:

Given a random variable X and a number $a \in \mathbb{R}$

$$(x-a)^2 = [(x-\mu) + (\mu-a)]^2 \quad (\text{where } \mu = E(x))$$

$$= (x-\mu)^2 + (\mu-a)^2 + 2(x-\mu)(\mu-a)$$

$$E(x-a)^2 = E(x-\mu)^2 + E(\mu-a)^2 + 2E[(\mu-a)(x-\mu)] \xrightarrow{} 0$$

$$E(X-a)^2 = var(x) + \underbrace{(\mu-a)^2}_{(\text{constant})}$$

→ CHEBYSHEV'S INEQUALITY:-

Let X be a non-negative r.v. with finite expectation. $(t>0)$ any positive real no.

Define a new r.v. Y from X

$Y = 0$ if $X < t$

$Y = t$ if $X \geq t$

$$E(Y) = 0 \cdot P(Y=0) + t \cdot P(X \geq t)$$

$$E(Y) = t \cdot P(X \geq t)$$

Note $X \geq Y$

$$E(X) \geq E(Y) = t \cdot P(X \geq t)$$

$$\Rightarrow \boxed{P(X \geq t) \leq \frac{E(X)}{t}} \quad — (*)$$

05/02/2019

→ Chebyshev's inequality -

X be a r.v. with mean μ and variance σ^2. For any real $t>0$,

$$\boxed{P(|X-\mu| > t) \leq \frac{\sigma^2}{t^2}}$$

Consider $(X-\mu)^2$.

$$P((X-\mu)^2 \geq t^2) \leq \frac{E(X-\mu)^2}{t^2} \quad (\text{From } \star)$$

★ Moment generating function :- (MGF)

$M_X(t) = E(e^{tX})$ for a given random variable X.

In case of a discrete r.v. X with p.m.f.

$$p(X = x_i) = f(x_i) \text{ for } x_i \in R_X,$$

$$M_X(t) = E(e^{tX})$$
$$= \sum_{x_i} e^{tx_i} p(X = x_i)$$

Ex:- $X \sim$ Bernoulli (p)

$$M_X(t) = e^{t\cdot 0}(1-p) + e^{t\cdot 1} p$$

$$M_X(t) = e^t p + (1-p).$$

Ex:- $X \sim$ Binomial (n, p) $\quad (q = 1-p)$

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (e^t p)^x (1-p)^{n-x}$$

$$= (q + pe^t)^n$$

Scanned by CamScanner

Ex: $X \sim \text{Poisson}(\lambda)$

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

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

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

Consider. $X \sim \text{Bernoulli}(p)$

$$M_X(t) = (1-p) + pe^t$$

$$\frac{d}{dt}(M_X(t)) = pe^t$$

$X \sim \text{Binomial}(n,p)$

$$M_X(t) = ((1-p) + pe^t)^n$$

$$\frac{d}{dt} M_X(t) = n((1-p)+pe^t)^{n-1} pe^t$$

$$(\text{at } t=0) = np \,!! \quad \text{How ??}$$

$$E(e^{tx}) = E\left(1 + \frac{tx}{1!} + \frac{t^2x^2}{2!} + \dots\right)$$

$$\cancel{\frac{d}{dt}E(e^{tx})} = E(1) + E(tx) + E\left(\frac{t^2x^2}{2}\right) + \dots$$

(absolutely summable)

$$= 1 + t\,E(x) + \frac{t^2}{2} E(x^2) + \dots$$

at $t=0$,

$E(1) = 1$

(absolutely convergent)

(thus d/dt and $+$ exchanged)

$$\frac{d}{dt}(E(e^{tx})) = 0 + 1\,E(x) + t\,E(x^2) + \frac{t^2}{2!} E(x^3) + \dots$$

Scanned by CamScanner

*** MAKE A TABLE - discrete, continuous for diff. distributions

at $t=0$,

$$\frac{d}{dt}\left(E(e^{tx})\right)\Big|_{t=0} = E(x)$$

$$\frac{d^2}{dt^2}\left(E(e^{tx})\right)\Big|_{t=0} = E(x^2)$$

⋮

(and so on).

★ Now Variance σ^2 can be derived from $E(x^2)$ and $E(x)$

(1) $M_X(t=0) = 1$; $M_X(t)$ is always defined for $t=0$.

(2) $M_X(t)$ as a funcn of t should be defined in a small interval around t.

(Q) Let X be a r.v. with MGF $M_X(t)$. Find the MGF of the r.v. $aX+b$ for some $a, b \in \mathbb{R}$, $a \neq 0$.

$$M_{aX+b}(t) = E\left(e^{(aX+b)t}\right)$$

$$= E\left(e^{atX} \cdot e^{bt}\right)$$

$$= e^{bt} \cdot E\left(e^{atX}\right)$$

$$M_{aX+b}(t) = e^{bt} M_X(at)$$

← Verify (conforms with pdf)