Probability & Stochastic Processes

$\Rightarrow$ r balls, n boxes

i) Both are distinct

$$n \times n \times n \times \cdots \times n = n^r$$

ii) Boxes are distinct, balls are not

$$\sum_{i=1}^{n} x_i = r \qquad \boxed{\text{Gen}} \to {}^{n+r-1}C_{n-1}$$

$$\boxed{\text{No empty Box}} \to {}^{r-1}C_{n-1}$$

$$ {}^{n}C_{r} = \underbrace{{}^{n-1}C_{r}}_{\text{(1) is not there}} + \underbrace{{}^{n-1}C_{r-1}}_{\text{(1) is there}}$$

iii) Balls (r) are distinct, Boxes (n) are not

$$S(r,n) = S(r-1,n-1) + n \cdot S(r-1,n)$$

$b_1, b_2, \cdots b_r$

a) b_1 is alone in a box $= S(r-1, n-1)$

b) b_1 with other elements $= n \cdot S(r-1,n)$

Boxes are non-empty

let K boxes are filled

$\Rightarrow$ $n-K$ boxes are empty

$\rightarrow S(r,K)$

$$\sum_{i=1}^{n} S(r,i)$$

Boxes can be Empty

iv) Balls are distinct & Boxes also distinct

No empty box $\boxed{n! \times S(r,n)}$

v) Both are identical

$$P(r,n) = P(r-1,n-1) + P(r-n,n)$$

$\hookrightarrow$ no empty allowed.

(a)

$P(8,3) = \begin{array}{l} 1,\boxed{1,6} \\ 1,\boxed{2,5} \\ 1,\boxed{3,4} \end{array} \to P(r-1,n-1) \to P(7,2)$

$\begin{array}{l} 2,\boxed{2,4} \\ 2,\boxed{3,3} \end{array} \to \boxed{\begin{array}{l} 1,1,3 \\ 1,2,2 \end{array}} \to P(r-n,n) \to P(5,3)$

$P(5,3) = P(4,2)$

* empty Box allowed $\to \sum_{i=1}^{n} P(r,i)$

$\to$ $f: X \to Y$

$|X| = r$ $(r>n)$ no of surjective function

$|Y| = n$

$= \boxed{n!\ S(r,n)}$

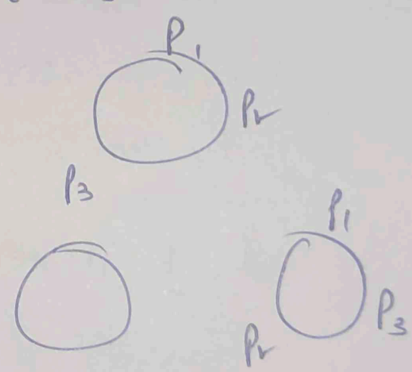
$\to$ Sterling number of 1st kind

Boxes turned into tables

$s(r,n)$

$P_1, P_2, \text{----}, P_r$

$\boxed{s(r,n) = s(r-1,n) + (r-1) \times s(r-1,n)}$



Cases$_r$

a) P_1 alone in a table : $S(r-1,n-1)$

b) P_1 with others : $S(r-1,n) \times (r-1)$

04/08/23

Probability and Stochastic processes

http://www.isical.ac.in/~arijit

↓

courses

↓

Probs & Stochastic Process [2023]

TA:

1. Debarshi Chanda
2. Uddalok Sarkar

Problems 1

$X = x_1, x_2, \dots x_i \dots x_n$ and x_i's are distinct real

goal: find the minimum / maximum number?

comparisions $= n-1$ $\boxed{\#\text{updates} \le \#\text{ comparisions}}$

$E[\#\text{updates}] = ?$

no of possibilities of updates for n number when the min occurs at ith position

$$= \frac{\binom{n}{i}(i-1)!\,(n-i)!}{n!} = 1/i$$

$$E[\#\text{updates}] \Rightarrow \sum_{i=1}^{n} 1/i \approx O(\log n)$$

Q) write a program, given n distinct numbers in an array, generate random permutation?

Backward Analysis:

P_i : Probability that min occurs at the ith position

$= 1/i$

$$E[\#\text{updates}] = \sum_{i=1}^{n} \left[\frac{1}{i}(1) + \left(1 - \frac{1}{i}\right)0\right]$$

Problem 2:

$X = \{x_1, x_2, \dots x_n\}$ and x_i's n distinct real numbers.

Find the median?

Answer 1:

Sort X and then pickup the "middle most" element

$$\left\lfloor \frac{n+1}{2} \right\rfloor$$

comparisions $= O(n \log n)$

- You have a biased coin that comes up Head (Success with a probability P.

Q) On the expectation how many times do you need toss a coin to get success? ↓ $\boxed{1/p}$

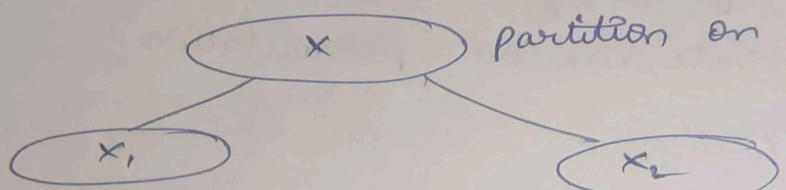
⇒ If we have a GP series with common-ratio (<1) asymptotically first dominates.

$\therefore n + \frac{n}{2} + \frac{n}{4} + \text{-----} = \boxed{n}$ ✓

* probability of picking the median is $\boxed{1/n}$
* probability of picking a "median-like" number is ?

Answer 2:

Can we do better?



algorithm:

i) pick a median-like number

ii) probability of picking the median = $1/n$

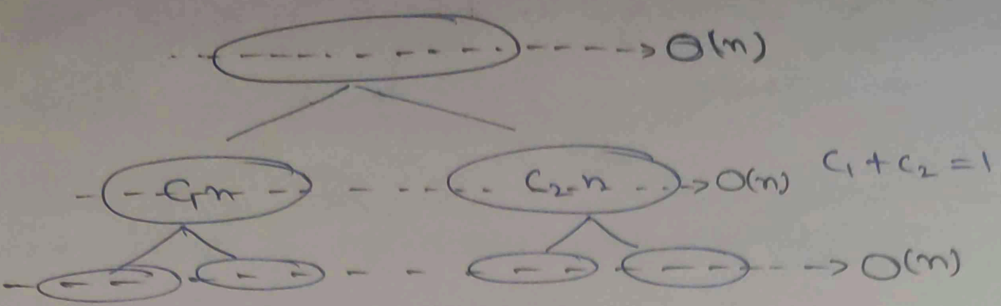
iii) probability of picking a median like number is constant

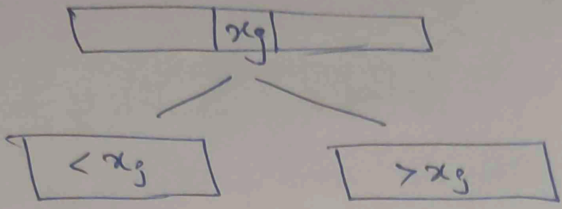
iv) Its like a amount of work is reducing

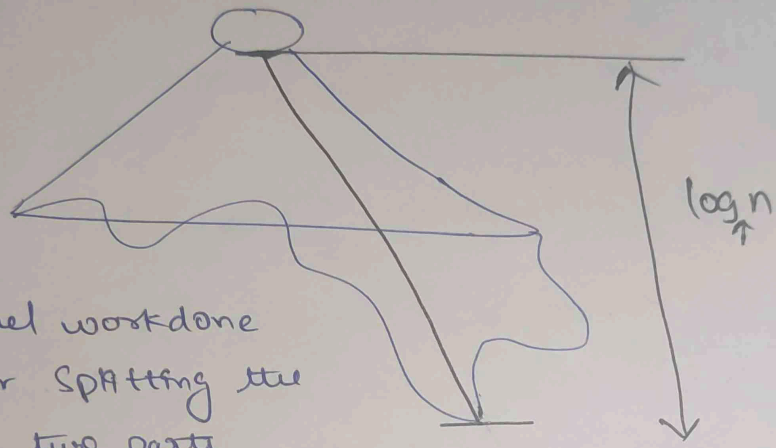
drawback:

It might never stop.

08/08/23



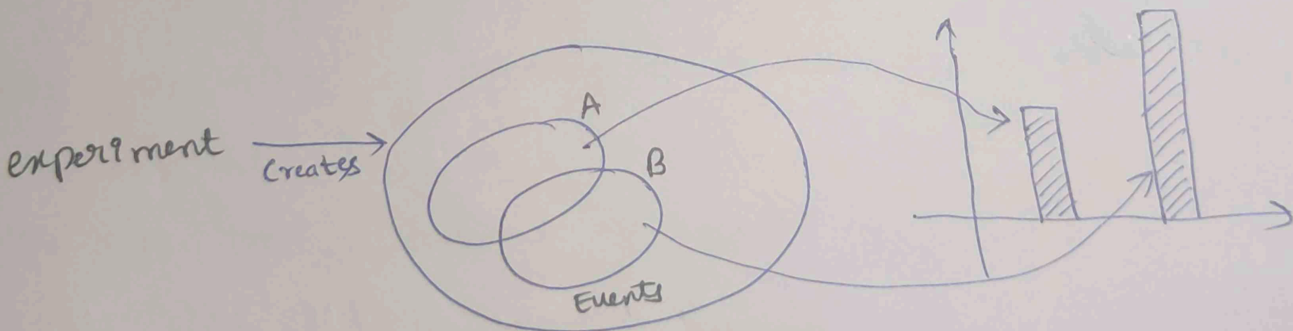




$\Rightarrow$ at each level workdone is $O(n)$ for splitting the array into two parts,

So, total time $= O(n \log_{\uparrow} n)$

but base may not be 2.



Sample Space (Ω)

(Set of all outcomes of the experiment)

Event : a Subset of Ω

let A, B be events

find $P(A \cup B)$, $P(A \cap B)$, $P(A \setminus B)$, $P(A \Delta B)$, ----

Defn: A collection F of subsets of Ω is called a σ-field if it satisfies the following

i) $\phi \in F$ and $\Omega \in F$,

ii) if $A \in F$ then $\bar{A} \in F$

and iii) if $A_1, A_2, \ldots \in F$ then $\bigcup_{i=1}^{\infty} A_i \in F$

exercise:

If $A_1, A_2, \ldots \in F$ then $\bigcap_{i=1}^{\infty} A_i \in F$

[to prove this we need ii & iii and De Morgan's law]

example:

$F = \{\phi, \Omega\}$, $F = \{\phi, A, \bar{A}, \Omega\}$

$F = \{0,1\}^{\Omega}$ this is power set (we can either take some element or leave it)

exercise:

Let $\{A_i : i \in I\}$ be a collection of sets. Prove that

(i) $\overline{\bigcup_i A_i} = \bigcap_i \bar{A_i}$ ii) $\overline{\bigcap_i A_i} = \bigcup_i \bar{A_i}$

exercise:

Let A and B belong to some σ-field F. Show that F contains the sets $A \cap B$, $A \setminus B$, $A \Delta B$.

notion of probability:

(Ω, F) is already defined

A probability measure on (Ω, F) is a real valued function

$$P: F \rightarrow [0,1]$$ satisfying the following

i) $P(\phi) = 0$, $P(\Omega) = 1$;

ii) if $A_1, A_2, \ldots$ is a collection of disjoint members of F, i.e., $A_i \cap A_j = \phi$ for all pairs i, j, $i \neq j$ then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

The triplet (Ω, F, P) comprising a set Ω, a σ-field F of subsets of Ω, and a prob. measure P on (Ω, F) is called a prob. space.

exercise:

Look at i) coin toss ii) roll of a dice.
and construct the probability spaces corresponding to them.

Probability distribution:

$$\{\omega_1, \omega_2, \omega_3, \ldots\}$$

$$P_1 + P_2 + P_3 + \cdots = 1$$

Some properties of (Ω, F, P):-

① $P(\bar{A}) = 1 - P(A)$

we know that $A \cup \bar{A} = \Omega$

A & $\bar{A}$ are disjoint

$\Rightarrow 1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$

② If $A \subseteq B$, then $P(A) \leq P(B)$ $\longrightarrow$ prove (use disjoint sets)

③ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$\Downarrow$$

$$P(A \cup B) \leq P(A) + P(B)$$

generalize

$$P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i)$$

[Union Bound / Boole's inequality]

Union Bound:

For any n ~~elements~~ events $A_1, A_2, \ldots A_n$

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

Proof: let $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_1 \cup A_2$, ----

----, $B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j$ ------, $B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j$

Note that B_i's are disjoint and

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

Also $B_i \subseteq A_i$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} P(B_i) \leq \sum_{i=1}^{n} P(A_i)$$

example:

consider the complete graph G on n vertices. Suppose, for an +ve integer k the following holds:

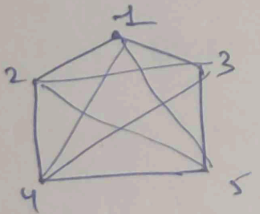
Then, it is possible to color the edge of these G red and blue so that no subgraph of k vertices has edges of just one color, i.e., (no monochramatic subgraph of size k).

Graph: (V,E)

V: Set of vertices

$|V| = n$ and $E \subseteq V \times V$

let $V = \{1, 2, 3, 4, 5\}$



In general we need to assume graph as a simple graph (no self loop, no parallel edges)

$$|E| \leq \binom{n}{2} = {}^nC_2$$

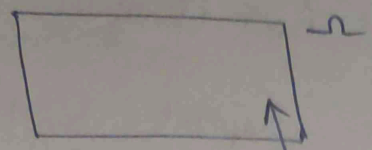
Subgraph is defaultly → induced sub graph

$G = (V, E)$ graph

$G' = (V', E')$ graph

$\boxed{V' \subseteq V}$ and all edges in G' should be b/w V'

Probabilistic method:



$P(\cdot) < 1$

$P(\cdot) > 0$ an element exists in space.

$P(\cdot) > 0$.

11/08/23

Continuity of probability measures:

The function P is a continuous set function.

This property says that P is countably additive rather than finitely additive

Theorem:

Let $A_1, A_2, \ldots$ be an increasing (decreasing) sequence of events

so that $A_1 \subseteq A_2 \subseteq \ldots$ ($A_1 \supseteq A_2 \supseteq \ldots$)

Define a limiting event

$$A = \bigcup_{i=1}^{\infty} A_i = \lim_{n\to\infty} A_n \quad \Big| \quad = \bigcap_{i=1}^{\infty} A_i = \lim_{n\to\infty} A_n$$

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

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

Proof: define $B_1 = A_1$, $B_2 = A_2 \setminus A_1$ ....

$$B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j = A_n \cap \overline{A_{n-1}}$$

B_n's $(n \geq 1)$ are disjoint $\bigcup_{i=1}^{n} B_i = A_n$

and also $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$

$$\Rightarrow P(A) = P\left(\bigcup_{i=1}^{\infty} A_i\right)$$
$$= P\left(\bigcup_{i=1}^{\infty} B_i\right)$$
$$= \sum_{i=1}^{\infty} P(B_i) \quad [\text{axiom}]$$
$$= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i)$$
$$= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} B_i\right)$$
$$= \lim_{n\to\infty} P(A_n)$$

example: Consider the complete graph on n vertices. Suppose for an +ve integer K the following holds: ?

Then it is possible to color the edges red and blue. So that no subgraph of K vertices has edges of just one color (no monochromatic subgraph of size K).

Proof:

color the edges uniformly at random and independent of one another,

→ In any subgraph of K vertices, the probability that every edge is red is $\left(\frac{1}{2}\right)^{\binom{K}{2}}$

→ There are $\binom{n}{K}$ subgraphs of size K. (# vertices)

→ Let A_i : event that the ith such subgraph is monochromatic.

$$\therefore \boxed{P(A_i) = 2 \times \left(\frac{1}{2}\right)^{\binom{K}{2}}}$$

$$P\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \le \sum_{i=1}^{\binom{n}{k}} P(A_i)$$

atleast one monochromatic subgraphs exists

$$= \binom{n}{k} 2 \cdot \left(\frac{1}{2}\right)^{\binom{k}{2}} < 1$$

$$\therefore P\left(\overline{\bigcup_{i=1}^{\binom{n}{k}} A_i}\right) > 0$$

probability of no subgraph is monochromatic

Principle of Inclusion and Exclusion:

If $A_1, A_2, \ldots, A_n$ are event then

$$\Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

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

Proof: use induction on n. Base case $(n=2)$

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

Derangement:

W.K.T. $[n] = \{1, 2, 3, \ldots n\}$

$$\pi : [n] \to [n]$$

$$\pi(i) \neq i \quad \forall i$$

ex) what is the probability that a permutation picked uniformly at random is a derangement?

Proof: let A_i : event that i occurs in its correct position in the permutation.

$\bigcup_{i=1}^{n} A_i \rightarrow$ atleast one number is in correct position

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

$$\cdots\cdots + (-1)^{n+1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

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

$$\cdots + (-1)^{n+1} \frac{1}{n!}$$

$$\simeq 1 - e^{-1}$$

$$\therefore \boxed{P\left(\overline{\bigcup_{i=1}^{n} A_i}\right) = e^{-1}}$$

Balls and Bins: (Birthday paradox)

→ Assume Birthdays are choosen independently and uniformly at random.

→ Let us count the # configurations where people (n) do not ~~some~~ share birthdays (365).

→ # possible configurations (with no restrictions) = 365^n

→ n persons can have n birthdays in n! ways and we have $\binom{365}{n}$ ways to select those birthdays.

→ So, the # configurations where people do not share birthdays = $\binom{365}{n} n!$

$$P(\text{no two persons share same bday}) = \frac{\binom{365}{n} n!}{365^n} \longrightarrow (1)$$

→ Viewed differently,

P(no two persons share the same bday)

$$= \left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right) \cdots \left(1-\frac{n-1}{365}\right) \longrightarrow (2)$$

→ Let us generalize to m persons and n birthdays

P(all m persons have different bdays)

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) \cdots \left(1-\frac{m-1}{n}\right)$$

$$= \prod_{i=1}^{m-1}\left(1-\frac{i}{n}\right)$$

$$\approx \prod_{i=1}^{m-1} e^{-i/n} \qquad \left[\because \text{ as } 1-\frac{x}{n} \approx e^{-x/n} \text{ when } x \ll n\right]$$

$$= \exp\left\{-\sum_{i=1}^{m-1}\frac{i}{n}\right\}$$

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

$$\approx e^{-m^2/2n}$$

Value of m at which this probability is 1/2 is approx

$$\frac{m^2}{2n} = \ln 2$$

$$\Rightarrow m = \sqrt{2n\ln 2}$$

when $n = 365$, $m = ?$

$$\cancel{m = \sqrt{\frac{2\times 365}{365^2}}} \qquad m = \sqrt{2\times 365\times \ln 2}$$

~~(Q)~~ Assume that $\sqrt{n}$ person has unique Bdays. Now find probability that next $\sqrt{n}$ person have unique Birthdays?

18/08/23

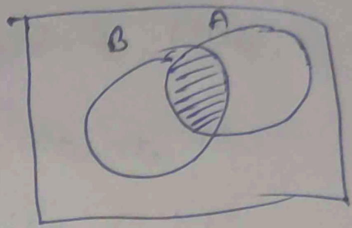
conditional probability :

ex) A family has two children. what is the prob that both are boys given that at least one is a boy? $\Rightarrow 1/3$ bcz here Sample Space $\Omega = \{GB, BG, BB\}$

definition :

If $P(B) > 0$, then the conditional probability that A occurs given that B occurs is defined to be

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



$\Rightarrow P(A/B) = C \cdot P(A \cap B)$

conditional probability of Ω given B must equal 1, and thus

$$C \cdot P(\cancel{\Omega \cap B}) = 1 \quad (\to B)$$

$$\Rightarrow \boxed{C = 1/P(B)}$$

Is conditional probability at all a probability measure?

i) $P(\Omega / B) = 1$

ii) For two disjoint events A_1 & A_2,

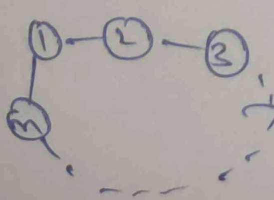
$$P(A_1 \cup A_2 / B) = P(A_1/B) + P(A_2/B)$$

min cut of Graph :

→ Its not unique

→ If we can get $V = V_1 \cup V_2$ and $V_1 \cap V_2 = \emptyset$ by "tearing away" minimum edges.

ex :



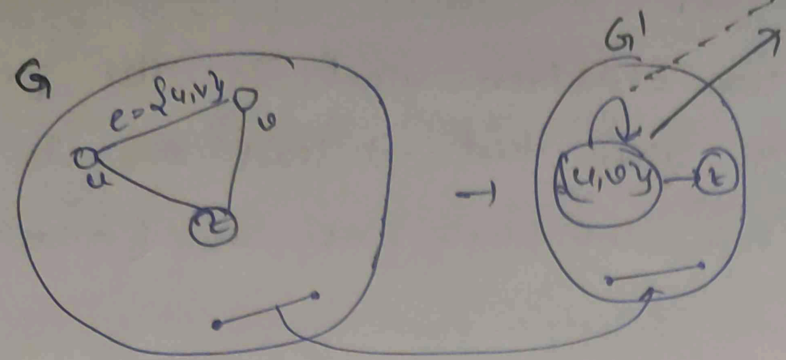
→ min-cut = 2

ex : (1)—(2)—(3)— - - - - —(n)

↓

min-cut = 1

Contraction of ~~edges/nodes~~ edges



Self loop should be there

Super Node | if uz and vz both exists in G

If we merge 2 edges in every step, till it is left (contracted) with 2 Super-Nodes.

1st iteration : I don't touch a min-cut edge

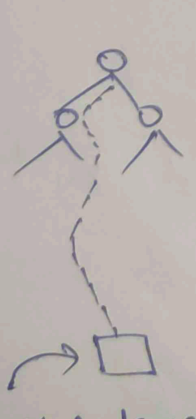
2nd iteration : I don't touch a min-cut edge | I didn't touch a min-cut edge in iteration 1.

3rd iteration : " " " " | " " " " iteration 1 and 2

(n-2)th iteration : " " " " | (... ∩ ...)

Modelling with conditional probability.

For problems having a sequential nature, first define conditional probability and then use them to determine unconditional probability.



The occurrence of a event is a sequence of traversals of the branches along the path from the root to the leaf.

an event of interest

we are dealing with an event A that occurs iff each one of the several events $A_1, A_2, \ldots, A_n$ has occurred i.e.,

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

Multiplication Rule:

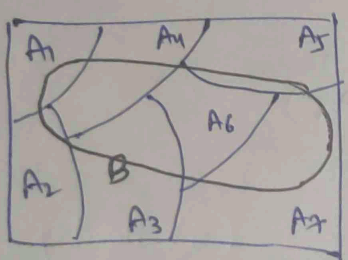
Assuming that all conditioning events have the probability, we have

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1).\, P(A_2/A_1).\, P(A_3/A_1 \cap A_2) \cdots \cdots P\left(A_n / \bigcap_{i=1}^{n-1} A_i\right)$$

Total Probability:

Let $A_1, A_2, \ldots A_n$ be disjoint events that form a partition of the sample space Ω and assume that

$P(A_i) > 0 \quad \forall \; i \in [n]$



Then for event B, we have

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)$$

$$= P(A_1)\, P(B/A_1) + P(A_2).\, P(B/A_2) + \cdots + P(A_n).\, P(B/A_n)$$

$$\therefore \boxed{P(B) = \sum_{i=1}^{n} P(A_i).\, P(B/A_i)}$$

ex)

	Prob of winning	Players
Gr I	0.3	50%
Gr II	0.4	25%
Gr III	0.5	25%

probability of winning?

Inference and Baye's rule:

The total probability theorem is used in conjunction with Baye's rule that relates conditional probability of the form $P(A/B)$ with conditional probability of the form $P(B/A)$, in which the ordering of the condition is reversed.

Baye's rule:

let $A_1, A_2, --- A_n$ be disjoint ~~elements~~ events that form a partition of the sample space Ω, and assume that

$$P(A_i) > 0 \ , \ \forall \ i \in [n],$$

~~of~~ Then for any event B, such that $P(B) > 0$, we have

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)}$$

$$= \frac{P(A_i) \cdot P(B/A_i)}{P(B)}$$

$$= \frac{P(A_i) \cdot P(B/A_i)}{\sum_{i=} P(A_i) \cdot P(B/A_i)}.$$

10) If a person has a disease test results are tue probability 0.9.

disease and test give -ve $= 0.9$

random person have disease drawn from the population $= 0.05$

given that the test has confirmed to have disease. what is the probability that the person actually have the disease?

20) Solve dearrangment using conditional probabilities in all possible arrangement, Probability that the picked arrangement is a dearangement?

30) N Rupees

x $N-x$ splitting between two persons.

Toss a coin, if head one wins, if tails other wins, find the probability that any one of them goes bankrupt?

22/08/23

Graph minimum cut :-

A Simple connected graph $G=(V,E)$ $|V|=n$ and $|E|=m$

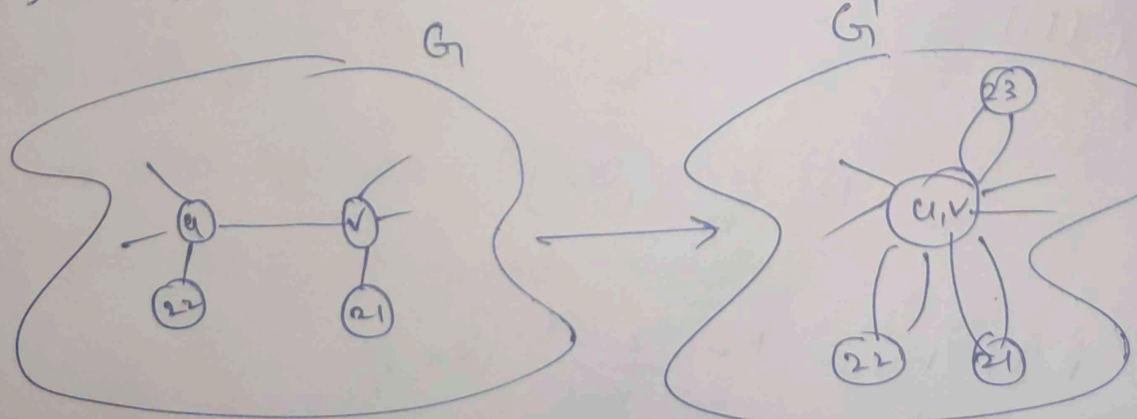
Let C be a mincut of r many min cuts $C_1, C_2, \ldots, C_r$

Algorithm :-

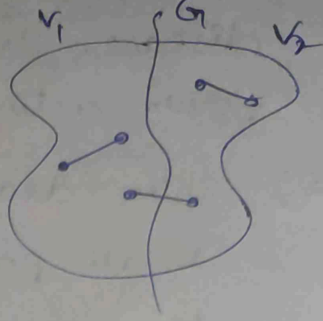
1) pick edges uniformly at random.
2) Contract edges. To generate a new graph.

\# iterations $= n-2$

3) Resultant graph has 2 super nodes and declare that as the min-cut.



E_i : event that in iteration i, no edge of C was "touched"



$$P\left(\bigcap_{i=1}^{n-2} E_i\right) = \ ? \quad (\text{Goal})$$

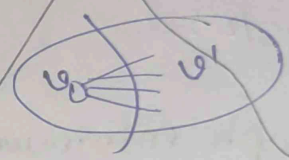
$$P(E_i) = ?$$

deg(v): degree of a vertex v in G

$$\sum_{v \in V} \deg(v) = 2|E|$$

It doesn't satisfy

~~Suppose $\exists$ a $v \in V$ such that $\deg(v) < K$~~



Lemma

$\forall v \in V$,

$$\deg(v) \geq K$$

$$\Rightarrow 2|E| = \sum_{v \in V} \deg(v) \geq n \cdot K$$

$$\Rightarrow |E| \geq \frac{nK}{2}$$

Let E_i: event that the Edge contracted in iteration i is not in C.

F_i: $\bigcap_{j=1}^{i-1} E_j$ event that no edge of C was contracted in the first. i iterations.

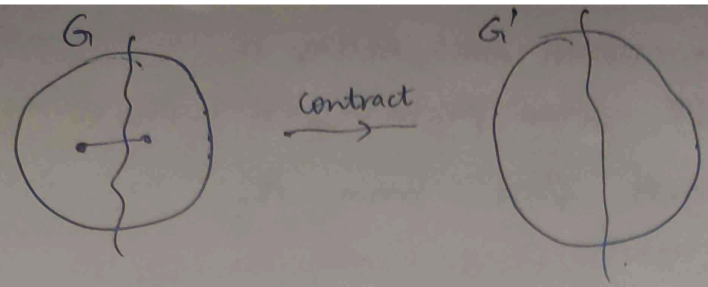
we need to compute $P(F_{n-2})$

$$P(E_1) \geq 1 - \frac{K}{nK/2}$$

$$1 - \frac{2}{n} \quad [\text{Ex}]$$

$$P(E_2 / F_1) \geq 1 - \frac{2}{n-1} \quad [\text{Ex}]$$

$\vdots$



Cuts are preserved under contraction.

$$P(E_i / F_{i-1}) \geq 1 - \frac{2}{n-i+1}$$

Now find $P(F_{n-2}) \geq \; ? \geq \frac{2}{n(n-1)} = \frac{1}{\binom{n}{2}}$

$$\leq 1 - \frac{1}{\binom{n}{2}}$$

$C_1, C_2, \text{----} \; C_r$

P(~~any one of the r cuts was formed~~ a particular cut was returned) $\geq \frac{1}{\binom{n}{2}}$

$C_1, C_2 \; \text{---} \; C_i \; \text{----} \; C_r \rightarrow$ min-cuts

Let E_i : event the cut C_i is returned

$$P(E_i) \geq \frac{1}{\binom{n}{2}}$$

$$\rightarrow 1 \geq P\left(\bigcup_{i=1}^{r} E_i\right) = \sum_{i=1}^{r} P(E_i) \geq \frac{r}{\binom{n}{2}}$$

$$\Rightarrow \boxed{r \leq \binom{n}{2}}$$

Gambler's Ruin

There are two players A and B. A has Rs. i and B has Rs. $N-i$. They toss a coin.

H (Success) with probability p

~~B~~ T (failure) with probability $1-p = q$

On success, A gets Rs. 1 from B.

On failure, A gives Rs. 1 to B.

The game is continued till anyone gets bankrupt. What is the probability that A end up with all the money?

Ans) E: event that A ends up with all the money (N) starting with Rs. i

let $P(E) = P_i$

Let us start by conditioning on first flip of the coin.

$$\Rightarrow P_i = P(E) = P(E/H) \cdot P(H) + P(E/\bar{H}) \cdot P(\bar{H})$$

$$\Rightarrow (p+q) P_i = p \cdot P(E/H) + q \, P(E/\bar{H})$$

(or)

$$p(P_{i+1} - P_i) = q(P_i - P_{i-1})$$

$$\boxed{\begin{array}{c} i = 1, 2, \ldots N-1 \\ p+q=1. \end{array}}$$

Boundary conditions, $P_0 = 0$, $P_N = 1$

$$P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1})$$

$$P_2 - P_1 = \frac{q}{p}(P_1 - \cancel{P_0}^{\,0}) = \frac{q}{p} P_1$$

$$P_3 - P_2 = \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 \cdot P_1$$

$$\vdots$$

$$P_i - P_{i-1} = \frac{q}{p}(P_{i-1} - P_{i-2}) = \left(\frac{q}{p}\right)^{i-1} P_1$$

$\vdots$

$$P_N - P_{N-1} = \frac{q}{p}\left(P_{N-1} - P_{N-2}\right) = \left(\frac{q}{p}\right)^{N-1} P_1$$

$$\Rightarrow P_i = \begin{cases} \dfrac{1 - (q/p)^i}{1 - (q/p)} P_1 & \text{if } p \neq q \\ i \cdot P_1 & \text{if } p = q = 1/2 \end{cases}$$

use the fact $P_N = 1$, to get

$$P_1 = \begin{cases} \dfrac{1 - q/p}{1 - (q/p)^N} & \text{if } p \neq 1/2 \\ 1/N & \text{if } p = 1/2 \end{cases}$$

$$\Rightarrow P_i = \begin{cases} \dfrac{1 - (q/p)^i}{1 - (q/p)^N} & \text{if } p \neq 1/2 \\ i/N & \text{if } p = 1/2 \end{cases}$$

Let Q_i: probability that B winds up with all when A starts with i and B with $N-i$.

$$Q_i = \begin{cases} \dfrac{1 - (p/q)^{N-i}}{1 - (p/q)^N} & \text{if } q \neq 1/2 \\ (N-i)/N & \text{if } q = 1/2 \end{cases}$$

verify what is

$$P_i + Q_i = ?$$

Q) we are given 3 coins and told that two of the coins are fair and the 3rd coin is biased, landing up H with probability 2/3.
we are not told which of the 3 coins is biased. We permute the coins randomly and then flip each of the coins. The 1st and 2nd coin come up H's and the 3rd coin is a T.
what is the prob that the first coin is a biased one?

$P(E_1)$ = Probability of 1st coin $= 1/3$

$P(E_2)$ = " " 2nd coin $= 1/3$

$P(E_3)$ = " " 3rd coin $= 1/3$

$P(E_1/B) = ?$ $\quad P(E_i) = 1/3 \quad \forall\, i = 1, 2, 3.$

$P(B/E_1) = ?$ $\quad \Rightarrow \dfrac{P(E_1 \cap B)}{P(B)}$

$P(B/E_2) = ?$

$P(B/E_3) = ?$ $\quad \Rightarrow \dfrac{P(B|E_1) \cdot P(E_1)}{P(B|E_1) \cdot P(E_1) + P(B|E_2) \cdot P(E_2) + P(B|E_3) P(E_3)}$

25/08/23

Q) Prove dearrangement using conditional probability.

E: event that no matches occur with n persons

$$\pi(i) \neq i \quad \forall\, i = [n]$$

M: event that a match occurs for the first person

$$P_n = P(E) = \overset{\to 0}{P(E|M)} \cdot P(M) + P(E|\bar{M}) \cdot P(\bar{M})$$

$$= P(E|\bar{M}) \left(\frac{n-1}{n}\right)$$

$$= P(E|\bar{M}) \left(1 - \frac{1}{n}\right)$$

C_n = # derangements for n persons

C_{n-1} 1 [2 ..... i ..... n]
1

$\pi(i) = 1$

Recurrence Relation would be

$$C_n = (n-1)(C_{n-2} + C_{n-1})$$

Base case $C_1 = 0$ and $C_2 = 1$

probability (this is using counting method)

$$\frac{C_n}{n!} = \frac{(n-1)}{n!}(C_{n-2} + C_{n-1})$$

$P(E/\bar{m}) = ?$

$$\boxed{P(E|\bar{m}) = P_{n-1} + \frac{P_{n-2}}{\cancel{n}-1}} \rightarrow E_x$$

Base cases $P_1 = 0$ and $P_2 = 1/2$ (b/t of 2 possibilities)

Independence

$P(A|B) = P(A)$ (or) $\frac{P(A \cap B)}{P(B)} = P(A)$ $\rightarrow$ $P(B) > 0$

$\Rightarrow P(A \cap B) = P(A) \cdot P(B)$

Q) Prove that disjoint events can never be independent.

$P(A \cap B) = P(\emptyset) = 0$

$\downarrow$ disjoint does not mean exhaustive

→ If there are n elements which are independent

$A_1 \; A_2 \; A_3 \; \cdots \; A_n$

what does Independent event mean?

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

useful for constructing Hash functions

all possible subsets of n elements.

Independence of collection of events:

we say that the events $A_1, A_2, \cdots, A_n$ are independent if

$$\boxed{P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)}$$

for every subset of $[n] = \{1, 2, \cdots, n\}$

ex) If A and B are independent, then

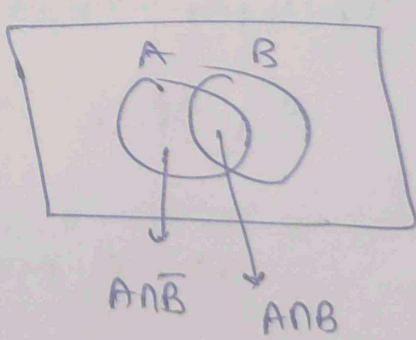
i) A and $\bar{B}$ are independent

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

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

proof

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



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

$\Rightarrow P(A \cap \bar{B}) = P(A) - P(A) \cdot P(B)$

$= P(A)\left[1 - P(B)\right]$

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

Note: $\boxed{\text{pairwise Independence} \not\Rightarrow \text{Independence}}$

* consider two independent fair coin tosses

H_1: { 1st toss is H } → HH, HT

H_2: { 2nd toss is H } → TH, HH

D: { the two tosses have diff results } → TH, HT

$$P(D) = 1/2$$

H_1 and H_2 are independent by definition.

What about D & H_1?

See, $P(D|H_1) = \dfrac{P(D \cap H_1)}{P(H_1)} = \dfrac{1/4}{1/2} = 1/2 = P(D)$ D & H_1 are indep

$P(D/H_2) = \dfrac{P(D \cap H_2)}{P(H_2)} = \dfrac{1/4}{1/2} = 1/2 = P(D)$ D & H_2 are indep

$$P(D \cap H_1 \cap H_2) = 0 \neq P(D).P(H_1).P(H_2)$$

* consider two independent rolls of a fair six-sided die and the following events:

A: { 1st roll is 1,2,3 } → $P(A) = 1/2$

B: { 1st roll is 3,4,5 } → $P(B) = 1/2$

C: { sum of two rolls is 9 } → $P(C) = 1/9$

$P(A \cap B \cap C) = \dfrac{1}{36} = P(A).P(B).P(C)$

$P(A \cap B) = \dfrac{1}{6} \neq P(A).P(B)$

$P(A \cap C) = \dfrac{1}{36} \neq P(A).P(C)$

$P(B \cap C) = \dfrac{1}{12} \neq P(B).P(C)$

$\boxed{\text{If K events are Independent} \not\Rightarrow \text{K-1 event are Indep}}$

Conditional Independence,

given an event C, the events A and B are conditionally independence if $P(C)>0$

$$P(A \cap B | C) = P(A|C) \cdot P(B|C) \longrightarrow (1)$$

for an alternate characterization,

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

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

$$P(A \cap B | C) = P(B|C) \cdot P(A | B \cap C) \longrightarrow (2)$$

compare (1) & (2)

$$P(A|C) \cdot P(B|C) = P(B|C) \cdot P(A | B \cap C).$$

$$\text{(or)} \quad P(A | B \cap C) = P(A|C) \longrightarrow (3)$$

So, conditional independence is equivalent to (3)

This states that if C is known to have occured, the additional knowledge that B also occured does not change the probability of A.

ex1) Two independent fair coin tosses, and all outcomes are

consider A to be 1st toss is Head

B to be 2nd toss is Head

C to have diff faces

Show that unconditional Indep does not imply conditional Indep

Q2) There are two coins Blue and Red, we choose one of two coins at random each with chosen probability $\frac{1}{2}$ and we proceed with two independent tosses. The coins are biased. with the blue coin the probability of Head is 0.9 and for read coin it is 0.1.

Define events as B = blue was selected

H_1 = 1st toss results in Head

H_2 = 2nd toss results in Head.

Notice the fact that given the toss H_1 and H_2 are independent

Show that conditional Independence does not implies unconditional Independence

29/08/23

Random Variables :-

A random variable X on a sample space Ω is a real valued function on Ω, i.e.,

$$X : \Omega \rightarrow \mathbb{R}.$$

A discrete random variable is a r.v. that takes only finite or countably infinite number of values.

"$X = x$"

X ↑ indicate r.v

x indicates value mapped to r.v.

$$\{ \omega \in \Omega \mid X(\omega) = x \}$$

$$\rightarrow \boxed{P(X=a) = \sum_{\omega \in \Omega : X(\omega)=a} P(\omega)}$$

definition:

Two r.v's are independent if and only if

$$P((X=x)\cap(Y=y)) = P(X=x).P(Y=y).$$

for all values of x, y.

Similarly we can define

R.v's $X_1, X_2, \ldots X_n$ are mutually independent if and only if for any subset $I \subseteq [n]$ and any value x_i, $i \in I$,

$$P\left(\bigcap_i X_i = x_i\right) = \prod_i P(X_i = x_i)$$

defn:

The expectation of a discrete R.V X denoted by $E[X]$, is given by

$$E[X] = \sum_i i.P(X=i)$$

where the sum is over all values in the range of X.

The expectation is finite if $\sum_i |i|.P(X=i)$ converges, otherwise the expectation is unbounded.

ex:

i) consider a r.v X that takes value 2^k with probability 2^{-k}, $k=1,2,3,\ldots$

ii) consider a r.v X that takes value 2^k and 2^{-k} with probability 2^{-k}. $k=2,3,\ldots$

Properties:

① For two r.v's, X and Y.

(linearity of expectation) $$E[X+Y] = E[X]+E[Y]$$

Proof: $E[X+Y] = \sum_i \sum_j (i+j)\, P((X=i) \cap (Y=j))$

→ Linearity of expectation holds for countably infinite summations in certain cases.

$$E\left[\sum_{i=1}^{\infty} X_i\right] = \sum_{i=1}^{\infty} E[X_i] \quad \text{whenever } \sum_{i=1}^{\infty} E[|X_i|] \text{ converges.}$$

* Note:
Linearity of expectation is not affected by (in)dependence of r.v's.

(2) For any constant c and a r.v X,

$$E[c.X] = c.E[X].$$

prop 1 continue → $E[X+Y] = \sum_i \sum_j i.\, P((X=i) \cap (Y=j)) + \sum_i \sum_j j.\, P((X=i) \cap (Y=j))$

$$= \sum_i i \sum_j P(X=i \cap Y=j) + \sum_j j \sum_i P(X=i \cap Y=j)$$

$$= \sum_i i\, P(X=i) + \sum_j j.\, P(Y=j)$$

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

prop 2 continue

Proof: $E[cX] = \sum_j j.\, P(cX=j)$

$$= c \sum_j \frac{j}{c}\, P(X=j/c)$$

$$= c \sum_k k.\, P(X=k)$$

$$= c.\, E[X].$$

Q) $E[X^2] \;\underline{\ ?\ }\; (E[X])^2$

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

By definition r.v. Y is non-ve

$$\Rightarrow E[Y] \geq 0.$$

$$E[Y] = E[X^2 - 2X.E[X] + E[X]^2]$$

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

$$\Rightarrow E[X^2] \geq E[X]^2$$

Jensen's Inequality:-

Shows that for any convex function f, we have

$$E[f(x)] \geq f[E[X])$$

A ~~constant~~ function $f: \mathbb{R} \rightarrow \mathbb{R}$ is said to be convex if for any x_1, x_2 and $0 \leq \lambda \leq 1$,

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

* **properties:-**

If f is twice differentiable, then f is convex if and only if $f''(x) \geq 0$.

Bernoulli trials:-

we have an experiment that Succeeds with Prob p and fails with a Prob $1-p$

$$X = \begin{cases} 1 & \text{if success} \\ 0 & \text{otherwise} \end{cases} \qquad P(X=1) = p$$

$$E[X] = P(X=1) = p$$

Binomial:-

we are performing n independent Trials each trial resulting in a success (p) or a failure (1-p).

Let x denotes a r.v indicating the # success

$$\boxed{P(X=j) = \binom{n}{j} p^j (1-p)^{n-j}}$$

find $E[X] = ?$

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

$X = X_1 + X_2 + \cdots + X_n$ (check!)

where each X_i is a Bernoulli trial with parameter p

$$\boxed{\therefore E(X) = \sum_{i=1}^{n} E[X_i] = np}$$

$\boxed{Var[X] = npq}$

Geometric r.v:

A geometric r.v X with parameter p is given by the probability distributions on $i = 1, 2, 3, \ldots$

$P(X=i) = (1-p)^{i-1} \cdot p$ || "waiting for success"

ex: Show that $\sum_{i \geq 1} P(X=i) = 1$

$E[X] = \sum_{i=0}^{\infty} i P(X=i)$

$= \sum_{i=1}^{\infty} i (1-p)^{i-1} p.$

$= \frac{p}{1-p} \sum_{i=1}^{\infty} i (1-p)^i$

ex1) memoryless guessing

ex2) with memory guessing

ex3) coupon collection problem

$\Rightarrow \boxed{E[X] = 1/p}$

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

$\Rightarrow \frac{1 \times q}{p^2} \Rightarrow \frac{1 \times q}{(1-(1-p))^2}$

ex 1) n cards with some number on it.

A machine w/o memory predicting numbers

what Expectation that ~~[illegible]~~ in the i^{th} iteration it don't chooses the already predicted numbers. Now, find Expected no of matches.?

$X_i \rightarrow$ R.V

for picking 1st card prob is $1/n$

if do for all n card $= n \times 1/n$

$= 1$

on i iteration, we have $n-(i+1)$ possibilities

ex) Coupon collector problem :

on i^{th} buy we should get new coupon to achieve the collection set, for getting offer.

what is Expected no of buys to get offer (collect all coupons (unique)).?

Best possible case is n, if we get all unique coupons in every i^{th} iteration.

$X_i \rightarrow$ #buys in phase i

$$\text{Success Probability} = \frac{n-(i+1)}{n}$$

$$\frac{1}{n}, \frac{1}{n-1}, \frac{1}{n-2}, \cdots, 1/2, 1$$

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{n}{n-i+1}$$

$$= n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} + 1\right)$$

$$\Rightarrow \boxed{E(X) = n \log n}$$

Here

ith phase

$(i-1)$ coupons I have — Go on until you get new coupon

x_i = No of buys in ith phase

$$E(x_i) = \frac{1}{p} = \frac{1}{\frac{n-(i-1)}{n}} = \frac{n}{n-i+1}$$

ex) There are n cards predict a number & match it with ith card. machine has memory. Find $E(x_i)$, x_i is a r.v. with ith match & $(i-1)$ is known.

Ans:

$$\frac{1}{n}, \frac{1}{n-1}, \frac{1}{n-2}, \cdots \frac{1}{2}, 1$$

$$\Rightarrow E(X) = \sum_{i=1}^{n} i . P(X=i)$$

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

$$= \sum_{i=1}^{n} \frac{1}{n-i+1}$$

$$E(x) = \log n$$

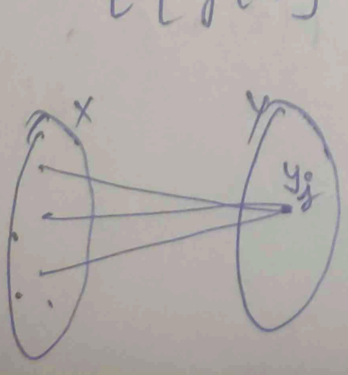
01/09/23

Functions of random variables:

Properties:- If X is a discrete r.v. that takes on one of the values x_i, $i \geq 1$, with respective probability $p(x_i)$, then for any real valued function g,

$$E[g(x)] = \sum_i g(x_i).p(x_i)$$

Proof:



$$\Rightarrow \sum_i g(x_i).\ p(x_i)$$

$$= \sum_j \sum_{i : g(x_i)=y_j} g(x_i).\ p(x_i).$$

$$= \sum_j y_j \sum_{i \cdot g(x_i) = y_j} p(x_i)$$

$$= \sum_j y_j \, P(g(x) = y_j)$$

ex: let X denote a r.v such that $X = \{-1, 0, 1\}$

with $P(X=-1) = 0.2$,

$P(X=0) = 0.5$,

$P(X=1) = 0.3$

find $E[X^2]$?

* If X and Y two independent r.v. then $f(X)$ and $g(Y)$ are also independent

pmf:

For a discrete r.v X, we define $p(a)$ of X by $p(a) = p(X=a)$. The pmf $p(a)$ is +ve for atmost a countable number of values of a. That is, ~~assume~~ if X must assume one of the values $x_1, x_2, ----$

then

$p(x_i) \geq 0$, $i = 1, 2, 3, --$

$p(x) = 0$ for all other x

|| Since X must take one of the values x_i we have

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

ex: The pmf of a r.v X is $P(i) = C \cdot \frac{\lambda^i}{i!}$, $i = 0, 1, 2, --$

where λ is +ve.

find $P(X=0)$.?

CDF :-

→ for a r.v. X the function $F(x) = P(X \leq x)$, $-\infty < x < \infty$ is called the CDF of X.

→ CDF specifies, for all real values of x, the prob that the r.v X is less than or equal to x.

$$F(a) = \sum_{\forall x \leq a} p(x)$$

→ CDF is a non-decreasing function of x.

Prop :-

$E[aX+b] = a\,E[X] + b$, a & b are constants. → H

memoryless property of a geometric r.v :-

Lemma :- for a geometric r.v. X with a parameter p and for $n \geq 0$, $P(X = n+k \mid X > k) = P(X = n)$

$$\Rightarrow P(X = n+k \mid X \geq k) = \frac{P(X = n+k \cap X \geq k)}{P(X \geq k)}$$

$$= \frac{P(X = n+k)}{P(X \geq k)}$$

$$= \frac{(1-p)^{n+k-1} \cdot p}{\sum_{i=k}^{\infty} (1-p)^{i-1} \cdot p}$$

$$= \frac{(1-p)^{n+k-1}}{(1-p)^k}$$

[since for $0 < x < 1$, $\sum_{i=k}^{\infty} x^i = \frac{x^k}{(1-x)}$]

$$= (1-p)^{n-1} \cdot p$$

$$= P(X=n).$$

Lemma II :- Let X be a discrete r.v. that takes only non -ve integer values.

Then $E[X] = \sum_{i=1}^{\infty} P(X \geq i)$

Proof :-
$$\sum_{i=1}^{\infty} P(X \geq i) = \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} P(X=j)$$

$$= \sum_{j=1}^{\infty} \sum_{i=1}^{j} P(X=j)$$

$$= \sum_{j=1}^{\infty} j \cdot P(X=j)$$

$$= E[X].$$

For a geometric r.v with parameter p

$$P(X \geq i) = \sum_{j=i}^{\infty} (1-p)^{j-1} \cdot p$$

Hence, $E[X] = \sum_{i=1}^{\infty} P(X \geq i)$

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

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

$$= 1/p$$

Markov's Inequality :-

Let X be a r.v, that takes only non-ve values, then for all $a>0$, $P(X \geq a) \leq$?

Proof: For $a>0$, let

$$I = \begin{cases} 1 & \text{if } X > a \\ 0 & \text{otherwise} \end{cases}$$

Notice that since X is non-ve, $(X \geq 0)$, $I \leq \frac{X}{a}$

But I is a 0-1 r.v.

$$\Rightarrow E[I] = P(I=1) = P(X \geq a)$$

$$\Rightarrow P(X>a) = P(I=1) = E[I] \leq E\left[\frac{X}{a}\right] = \frac{E[X]}{a}$$

ex) n independent tosses of a fair coin, $P(X \geq 3n/4)$

Let $X_i = \begin{cases} 1 & \text{if the } i^{th} \text{ toss gives a H} \\ 0 & \text{otherwise} \end{cases}$ $\qquad X = \sum_{i=1}^{n} X_i$

then we can write it as

$P(X_i = 1) = 1/2$

$E[X_i] = P(X_i = 1) = 1/2$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = n/2$$

$$P(X \geq 3n/4) \leq \frac{E[X]}{3n/4} = \frac{n/2}{3n/4} = 2/3$$

__Variance and moments of a r.v :-__

defn:- The k^{th} moment of a r.v is $E[X^k]$

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

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

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

$$\therefore Var[X] = \underbrace{E[X^2]}_{\text{2nd moment}} - \underbrace{(E[X])^2}_{\text{1st moment}}$$

$\therefore$ Standard deviation $\sigma[X] = \sqrt{Var[X]}$

defn: The covariance of two r.v's X and Y is

$$\boxed{E[(X-E[X])(Y-E[Y])] = E(x\cdot y) - E(x)\cdot E(y)}$$

lemma: For any two r.v's X and Y,

$$\boxed{Var[X+Y] = var[X] + var[Y] + 2\cdot cov(X,Y)}$$

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

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

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

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

$\therefore Var[X+Y] = Var[X] + var[Y] + 2\cdot cov(X,Y)$

ex) for any finite collection of r.v's $X_1, X_2, \cdots X_n$

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} var[X_i] + 2\sum_{i=1}^{n}\sum_{j=i}^{n} cov(X_i, X_j)$$

Lemma:

If X and Y are two independent r.v's then

$E[X\cdot Y] = E[X]\cdot E[Y]$ (viceversa is not true)

$$\left\{\begin{array}{l} X \,\&\, Y \text{ ind r.v's} \Rightarrow E[X\cdot Y] = E[X]\cdot E[Y] \\ E[XY] = E(X)\cdot E(Y) \not\Rightarrow X \,\&\, Y \text{ ind r.v's} \end{array}\right\}$$

proof:

$$E[X.Y] = \sum_i \sum_j (i.j)\, P((X=i) \cap (Y=j))$$

$$= \sum_i \sum_j (i.j).\, P(X=i).\, P(Y=j)$$

$$= \left(\sum_i i\, P(X=i).\right) \left(\sum_j j\, P(Y=j)\right)$$

$$= E[X].\, E[Y].$$

Corollary:

For two independent r.v's X and y, then $\operatorname{cov}(X,Y) = 0$

Proof: Ex.

Theorem:

(pairwise independence)

$\Uparrow$

If $X_1, X_2, \ldots X_n$ are "mutually ind" r.v's then

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i].$$

05/09/23

Mean and variance of some r.v's:-

① Bernoulli Trial:

$X \sim$ Bernoulli (P)

mean $= P$ and variance $= P(1-P)$

② Binomial:

$X \sim$ Binomial (n, P)

mean $= nP$ and variance $= nP(1-P)$

③ Poisson:

$X \sim$ Poisson (λ)

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

where $i = 0, 1, 2, \ldots$

λ is the

ex: i) $\boxed{\text{mean} = \lambda}$ and $\boxed{\text{variance} = \lambda}$

ii) check that it is a pmf

$$\left.\begin{array}{l} n \to \text{large} \\ p \to \text{small} \end{array}\right\} \text{then} \quad np = \lambda \quad \to \text{(I)}$$

where λ is some finite value

Proof:

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

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \frac{n!}{(n-i)!\,(i)!} p^i (1-p)^{n-i}$$

$$= \frac{n(n-1)\cdots(n-i+1)}{i!} p^i (1-p)^{n-i}$$

from (I) $\boxed{p = \frac{\lambda}{n}}$

$$= \frac{n(n-1)\cdots(n-i+1)}{i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$$

$$= \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \left(1 - \frac{\lambda}{n}\right)^{n-i}$$

for a fixed i, let $n \to \infty$

for $j = 1, 2, \ldots, i$, we have

$$\frac{n-i+j}{n} \to 1, \quad \left(1 - \frac{\lambda}{n}\right)^{-i} \to 1$$

Then for fixed i the above exp tends to

pmf(λ)

(4) Geometric :-

$X \sim Geometric(P)$

$\boxed{mean = 1/p}$ and $\boxed{Variance = \frac{1-p}{p^2}}$

(5) discrete uniform :-

$X \sim Uniform[a, b]$

$$P(X = i) = \begin{cases} \frac{1}{b-a+1} & , i = a, a+1, \ldots, b \\ 0 & , \text{otherwise} \end{cases}$$

$\boxed{mean = \frac{(a+b)}{2}}$ and $\boxed{Variance = \frac{(b-a)(b-a+2)}{12}}$

* If we shift origin we know that

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

$X_1 \sim U[a, b]$ and $X_2 \sim U[1, b-a+1]$

$Var[X_1] = Var[X_2]$

Easy to calculate $Var[X]$ when $X \sim U[1, n]$

$$Var[X] = \frac{n^2-1}{12}$$

Mean and Variance of Sample mean :-

iid : independent and identically distribution

Let $X_1, X_2, \ldots X_n$ be iid r.v's

Let $E[X_i] = \mu \quad \forall i = 1(1)n$ 1 to n

$Var[X_i] = \sigma^2 \quad \forall i = 1(1)n$

Let $X = \frac{\sum_{i=1}^{n} X_i}{n}$

$$E[\bar{X}] = \mu$$

$$\Rightarrow var[\bar{X}] = var\left[\frac{\sum X_i}{n}\right] = \frac{1}{n^2} var[\sum X_i]$$

$$= \frac{\sum var[X_i]}{n^2}$$

$$= \frac{n\sigma^2}{n^2}$$

$$= \sigma^2/n$$

ex: Approval rating.

$\boxed{\text{exercise: find the implication for the random variable } X \text{ when } var(X)=0.}$

Chebyshev's Inequality:

For any $a>0$, $P(|X-E[X]| \geq a) \leq \frac{var[X]}{a^2}$

Proof:-

First observe that $P(|X-E[X]| \geq a) = P((X-E[X])^2 \geq a^2)$

Since $(X-E[X])^2$ is a non -ve r.v., we can apply Markov's Inequality to get

$$P(|X-E[X]| \geq a) = P((X-E[X])^2 \geq a^2)$$

$$\leq \frac{E[(X-E[X])^2]}{a^2}$$

$$\leq \frac{var[X]}{a^2}$$

Corollary: For any $t>1$, $P(|X-E[X]| \geq t.\sigma[X]) \leq \frac{1}{t^2}$

and $P(|X - E[X]| \geq t \cdot E[X]) \leq \frac{var[X]}{t^2 \cdot (E[X])^2}$

ex) n ind fair coin tosses.

X : # heads.

$P(X \geq \frac{3n}{4}) \leq ?$

check $E[X] = ?$ and $var[X] = ?$

X, we can write as sum of independent Bernoulli's

$X = \sum_{i=1}^{n} X_i$ $\quad X_i = \begin{cases} 1 & \text{if } i\text{th toss is } H \\ 0 & \text{otherwise} \end{cases}$

$P(X_i = 1) = \frac{1}{2} = E[X_i]$

$var[X_i] = 1/4$?

So, $E[X] = \frac{n}{2}$ and $var[X] = n/4$

$$\Rightarrow P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{3n}{4} - \frac{n}{2})$$

$$= P(X - E[X] \geq \frac{n}{4})$$

$$\leq P(|X - E[X]| \geq \frac{n}{4})$$

$$\leq \frac{var[X]}{(n/4)^2} = \frac{(n/4)}{(n/4)^2} = 4/n$$

ex) Let X be a r.v taking values in the range $[a, b]$.

Then, $var[X] \leq (b-a)^2/4$

ex) Let p be the fraction of voters who support the party. we interview n voters chosen ~~~~ independently and uniformly at random for population.

So the reply of voter can be modelled as independent Bernoulli with the parameter p.

$X_i \sim$ Bernoulli (P), there are n ind. Bernoulli trials

Let $M_n = \frac{\Sigma X_i}{n}$, we return M_n as our estimate

$$E[M_n] = P. \quad var[M_n] = \frac{P(1-P)}{n}$$

we are interested in how far M_n is from p.

Chebyshev's tells us [luckily $E[M_n] = P$!]

$$P(|M_n - \overset{E[M_n]}{P}| \geq \varepsilon) \leq \frac{var[M_n]}{\varepsilon^2}$$

$$= \frac{P(1-P)}{n\varepsilon^2}$$

$$= \frac{var[X_i]}{n\varepsilon^2}$$

$$\leq \frac{1}{4n\varepsilon^2} \quad \text{(bcz of above ex)}$$

let $\varepsilon = 0.01$

Prob of error ≤ 0.05

08/09/23

Joint PMF's of multiple r.v's:

→ All such r.v's associated with the same experiment, Same sample space and prob law

Joint pmf is $P((X=x) \cap (Y=y)) = P(X=x, Y=y)$

If A is the set of all pairs (x,y) that have a certain property then

$$P((X,Y) \in A) = \sum_{(x,y) \in A} P(X=x, Y=y)$$

Note that I

i) marginal pmf of x

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

ii) marginal pmf of y

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

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

$$= \sum_{y} pmf_{X,Y}(x,y)$$

<u>functions of r.v</u>:

A function $Z = f(X,Y)$ of r.v's X and Y defines another r.v

$$P(Z=z) = \sum_{\{(x,y) \mid f(x,y)=z\}} P(X=x, Y=y)$$

$$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y)\, pmf_{X,Y}(x,y)$$

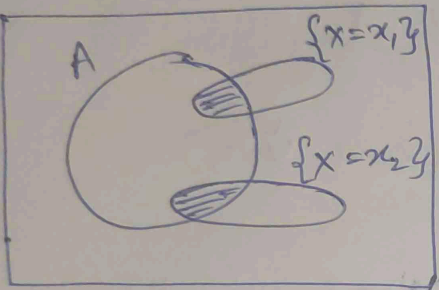
<u>ex</u>:- In the derangement problem, what is the expected # people who get back their own umbrella?

Conditioning: (conditional pmf)

Conditioning a r.v on an event: The conditional pmf of a r.v X, conditioned on a particular event A, with $P(A) > 0$, is defined by

$$\text{pmf}_{X|A}(x) = P(\{X=x \mid A\}) = \frac{P(\{X=x\} \cap A)}{P(A)} \rightarrow ①$$

→ Events $\{X=x\} \cap A$ are disjoint for diff^n values of x and their union is A.



So, $P(A) = \sum_x P(\{X=x\} \cap A) \rightarrow ②$

from ① & ②

$$\Rightarrow \boxed{\sum_x \text{pmf}_{X|A}(x) = 1}$$

ex) Let X be the roll of a fair die and A: event that the roll is even

$$P(X=k \mid A) = \frac{P(\{X=k\} \cap A)}{\sum_x P(\{X=x\} \cap A)}$$

ex) A student takes a test Repeatedly, upto a max of n times, each time with a prob p of passing, independent of the # of previous attempts. What is the pmf of the # attempts, given that the student passes the test?

Let A: event that the student passes the test with atmost n attempts

$A = \{X \leq n\}$. $P(A) = \sum_{i=1}^{n} (1-p)^{i-1} p.$

$$pmf_{X|A}(k) = \begin{cases} \dfrac{(1-p)^{k-1} \cdot p}{P(A)} & , \text{if } k = 1, 2, \ldots n \\ 0 & , \text{otherwise} \end{cases}$$

<u>Conditioning a r.v on another r.v</u>:

The conditional $pmf_{X|Y}$ of X given Y, is defined by specializing the previous definition of $pmf_{X|A}$ to events A of the form $\{Y = y\}$

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

use the definition of conditional prob, to get

$$pmf_{X|Y}(x/y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{pmf_{X,Y}(x, y)}{pmf_Y(y)}$$

ex: fix some y with $pmf_Y(y) > 0$ and consider $pmf_{X|Y}(x|y)$ as a function of x, verify that this function is a valid pmf of X.

$$pmf_{X,Y}(x, y) = pmf_Y(y) \cdot pmf_{X|Y}(x|y)$$
$$= pmf_X(x) \cdot pmf_{Y/X}(y/x)$$

The conditional pmf can also be used to calcula the marginal pmf.

$$pmf_X(x) = \sum_y pmf_{X,Y}(x,y)$$

$$= \sum_y pmf_Y(y) \cdot pmf_{X|Y}(x|y)$$

↳ "divide and conquer"

→ If $A_1, A_2, \ldots A_n$ are 'disjoint events' and forms a partition of Ω, with $P(A_i) > 0$ $\forall i$ then

$$\boxed{pmf_X(x) = \sum_{i=1}^{n} P(A_i) \cdot pmf_{X|A_i}(x).}$$

proof: $pmf_X(x) = P(X=x)$

$$= P(\{X=x\} \cap A_1) + \ldots + P(\{X=x\} \cap A_n)$$

$$= P(A_1) \cdot P(\{X=x\}|A_1) + \ldots + P(A_n) \cdot P(\{X=x\}|A_n)$$

$$= \sum_{i=1}^{n} P(A_i) \cdot pmf_{X|A_i}(x).$$

→ for any event B with $P(A_i \cap B) > 0$ $\forall i$,

$$pmf_{X|B}(x) = \sum_{i=1}^{n} P(A_i \cap B) \cdot pmf_{X|A_i \cap B}(x)$$

Conditional expectation:

Let X and Y r.v's with the same experiment. The conditional exp of X given an event A with $P(A) > 0$, is defined by

$$\boxed{E[X|A] = \sum_x x \cdot pmf_{X|A}(x)}$$

* For a function $f(x)$, we have

$$E[f(x)|A] = \sum_x f(x) \cdot pmf_{X|A}(x)$$

→ The conditional expectation of X given a value of Y is defined by

$$E[X \mid Y=y] = \sum_{x} x \cdot pmf_{X|Y}(x|y)$$

→ If $A_1, A_2, \ldots, A_n$ be disjoint events that form a partition of Ω, with $P(A_i) > 0 \ \forall i$,

then $$E[X] = \sum_{i=1}^{n} P(A_i) \cdot E[X \mid A_i]$$

Proof:- $$E[X] = \sum_{x} x \cdot pmf_X(x)$$

$$= \sum_{x} x \cdot \sum_{i=1}^{n} P(A_i) \cdot pmf_{X|A_i}(x)$$

$$= \sum_{i=1}^{n} P(A_i) \cdot \underbrace{\sum_{x} x \cdot pmf_{X|A_i}(x)}$$

$$= \sum_{i=1}^{n} P(A_i) \, E[X \mid A_i]$$

→ for any event B with $P(A_i \cap B) > 0$, $\forall i$ we have

$$E[X \mid B] = \sum_{i=1}^{n} P(A_i \mid B) \cdot E[X \mid A_i \cap B]$$

$$E[X] = \sum_{y} pmf_Y(y) \cdot E[X \mid Y=y]$$

defn:
The expression $E[X \mid Y]$ is a r.v $f(Y)$ that takes on the values $E[X \mid Y=y]$ when $Y = y$

$$E[X \mid Y] : \Omega \to \mathbb{R}$$

ex: Ind roll two standard dice. Let X_1 : # on the 1^{st} and X_2 : # on the 2^{nd}.

X : Sum of the # on both.

$$E[X \mid X_1] = \sum_x x \cdot P(X = x \mid X_1)$$

$$= \sum_{x = X_1 + 1}^{X_1 + 6} x \cdot \frac{1}{6}$$

$$= X_1 + \frac{7}{2}$$

Now find $E[X]$?

Note: $E[E[X \mid Y_1]] = E[X]$

Do it in the usual brute force way, you will get the same answer

Theorm: $E[X] = E[E[X \mid Y]]$

Proof: $E[E[X \mid Y]] = \sum_y E[X \mid Y = y] \cdot P(Y = y)$

$= E[X]$ {by the total expectation theorem. $E[X] = \sum_y E[X \mid Y = y] \, pmf_Y(y)$ }

Linearity of conditional Expectation:

For any finite collection of discrete r.v's $X_1, X_2, \ldots X_n$ with finite exp and for any r.v Y,

$$E\left[\sum_{i=1}^{n} X_i \mid Y = y\right] = \sum_{i=1}^{n} E[X_i \mid Y = y]$$

* find mean and variance Geometric r.v?

12/09/23

Expectation and variance of a geometric r.v : (Parameter: p)

$X \sim Geo(p)$

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

$P(X=1) = p$

where $x = 1, 2, 3, \ldots\ldots$

$P(X>1) = 1-p$

Let $A_1 = \{X=1\}$: first try is success

$A_2 = \{X>1\}$: first try is failure

If the first try is successful, $X=1$ and $E[X|X=1]=1$

If the first try fails, ie if A_2 has occured, then we have wasted a try and we are back to where we started.

So, the expected # remaining tries is $E[X]$, and we get

$$E[X|X>1] = E[X+1] = E[X]+1$$

use total expectation theorem

$$E[X] = P(X=1) \cdot E[X|X=1] + P(X>1) \cdot E[X|X>1]$$

$$\Rightarrow E[X] = p \cdot (1) + (1-p)(E[X]+1) \Rightarrow \boxed{E[X] = 1/p}$$

Variance:

$E[X^2|X=1] = 1$, $E[X^2|X>1] = 1+2E[X]+E[X^2]$

$$\Rightarrow E[X^2] = P(X=1) \cdot E[X^2|X=1] + P(X>1) \cdot E[X^2|X>1]$$

ex) $X_1, X_2, \dots X_N$ iid r.v's where N is a r.v taking the integer values

$N, X_1, X_2, \dots \rightarrow$ independent

$Y = X_1 + X_2 + \dots + X_N$

$E[Y] = ?$ conditioned on the fact that $N=n$

$= E[N] \cdot E[X_1]$

$= n \cdot E(X_1]$

Branching process:

Consider a program that includes one call to a process S. Assume that each call to S recursively spawns new copies of the process S. where the # new copies $\sim B(n, p)$. we assume that these r.v's are ind for each call of S. what is the expected # copies of S generated?

Soln:-

Introduce the notation of generations. The initial process is in generation 0. O/w a process is in generation i if it was spawned by a process in generation $i-1$,

Let Y_i : #process in generation i.

$Y_0 = 1 \rightarrow ①$ $\quad Y_1 \sim Bin(n, p)$

$E[Y_1] = np. \rightarrow ②$ Suppose, # process in generation $i-1$ is y_{i-1}

So, $Y_{i-1} = y_{i-1}$. Let Z_K : # copies spawned by the K^{th} process that was spawned in the ~~i^{th}~~ $(i-1)^{th}$ generation.

$K = 1, 2, \ldots, y_{i-1}$

Each Z_K is a binomial r.v with parameter n and p.

Then $$E[Y_i \mid Y_{i-1} = y_{i-1}] = E\left[\sum_{k=1}^{y_{i-1}} Z_K \mid Y_{i-1} = y_{i-1}\right]$$

$$= \sum_{j \geq 0} j \cdot P\left(\sum_{k=1}^{y_{i-1}} Z_K = j \mid Y_{i-1} = y_{i-1}\right)$$

As Z_k's are all ind binomial r.v's

$$= \sum_{j \geq 0} j \cdot P\left(\sum_{k=1}^{y_{i-1}} Z_K = j\right)$$

$$= E\left[\sum_{k=1}^{y_{i-1}} Z_k\right] = \sum_{k=1}^{y_{i-1}} E[Z_K] = y_{i-1} \cdot np \longrightarrow (3)$$

Now, $$E[Y_i] = E[E[Y_i \mid Y_{i-1}]]$$

$$= E[Y_{i-1} \cdot np]$$

$$= np \cdot E[Y_{i-1}] \longrightarrow (4)$$

By using induction on i, and using the fact $Y_0 = 1$,

we have $$E[Y_i] = (np)^i$$

The expected #copies of the process S generated by the program is $$E\left(\sum_{i \geq 0} Y_i\right) = \sum_{i \geq 0} E[Y_i] = \sum_{i \geq 0} (np)^i$$

CLASS - 14 Dt. 26/9/23

Moment Generating Function (MGF) :- The MGF of a random variable X is defined as $M_X(t) = E[e^{tX}]$

The function $M_X(t)$ captures all moments of the random variable X ;

Theorem: Let X be a random variable with MGF $M_X(t)$. We assume that exchanging the expectation and differentiation operands is legitimate ; [**Note**: We can do this when the MGF exists around in the neighbourhood of 0] , then $\forall n > 1$, $E[X^n] = M_X^{(n)}(0)$ where, $M_X^{(n)}(0)$ is the nth derivative of $M_X(t)\big|_{t=0}$;

Proof: Because of the assumption,

$$M_X^{(n)}(t)\Big|_{t=0} = E[X^n \cdot e^{tX}]\Big|_{t=0} = E[X^n]$$

ex: Deduce the mgf of a geo. r.v. $X \sim Geo(p)$

For, $t < -\ln(1-p)$

$$M_X(t) = E[e^{tX}] = \sum_{x=1}^{\infty} (1-p)^{x-1} \cdot p \cdot e^{tx}$$

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

$$M_X^{(1)}(t)\Big|_{t=0} = ? = \frac{p}{1-p}\left[\left(1-(1-p)e^t\right)^{-2} \cdot \left(\cancel{-}(1-p)\right)\right]\Big|_{t=0}$$

$$= \frac{p}{1-p} \cdot \frac{1-p}{(\cancel{1}-\cancel{1}+p)^2} = \frac{1}{p}$$ @.

Th: If X and Y are ind. r.v.s , then $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

Proof: $M_{X+Y}(t) = E\left[e^{t(X+Y)}\right]$

$$= E[e^{tX} \cdot e^{tY}] = E[e^{tX}] \cdot E[e^{tY}]$$

$$= M_X(t) \cdot M_Y(t)$$

Deriving and Applying Chernoff Bounds :- The chernoff bounds for a r.v. X is obtained by applying Markov's inequality to e^{tX} for some well chosen value of t ;

From Markov's inequality, for any $t > 0$,

$$P(X \geq a) = P(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{at}}$$

In particular, $P(X \geq a) \leq \min_{t>0} \left(\frac{E[e^{tX}]}{e^{at}} \right)$

lly for $t<0$, $P(X \leq a) = P(e^{tX} \geq e^{at}) \leq \frac{E[e^{tX}]}{e^{at}}$;

$$P(X \leq a) \leq \min_{t<0} \left(\frac{E[e^{tX}]}{e^{at}} \right)$$

Chernoff for Sum of (ind.) Poisson Trials:-

We are interested in tail inequalities for the sum of 0-1 independent r.v.s (known as poisson trials)

Let, $X_1, X_2, X_3, \cdots, X_n$ be n ind. Poisson Trials s.t., $P(X_i = 1) = p_i$. Let, $X = \sum_{i=1}^{n} X_i$. Then $E[X] = \sum_{i=1}^{n} p_i = \mu$

To compute the Chernoff bounds, we need to compute the MGF of X, which in turn needs us to compute the MGF of X_is;

$$M_{X_i}(t) = E[e^{tX_i}] = p_i e^t + (1-p_i) = 1 + p_i(e^t - 1) \leq e^{p_i(e^t-1)}$$

$[\because \forall y \quad 1+y \leq e^y]$

$$M_X(t) = E[e^{tX}] = E[e^{t(X_1+X_2+\cdots+X_n)}]$$

$$= \prod_{i=1}^{n} E[e^{tX_i}]$$

$$\leq \prod_{i=1}^{n} \left(e^{p_i(e^t-1)}\right)$$

$$= \exp\left\{\sum_{i=1}^{n} p_i(e^t-1)\right\} = e^{\mu(e^t-1)}$$

$\therefore M_X(t) \leq e^{\mu(e^t-1)}$

Th: Let $X_1, X_2, \cdots, X_n$ be ind. poisson trials, s.t. $P(X_i = 1) = p_i$ and $X = \sum_{i=1}^{n} X_i$. Let, $E[X] = \mu$. Then, the following holds:-

(i) For any $\delta > 0$, $P(X \geq (1+\delta)\mu) \leq \left\{ \frac{e^\delta}{(1+\delta)^{(1+\delta)}} \right\}^\mu$.

(ii) For any $0 < \delta < 1$, $P(X \leq (1-\delta)\mu) \leq \left\{ \frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}} \right\}^\mu$

Proof:- Apply Markov's inequality for any $t > 0$,

$$P(X \geq (1+\delta)\mu) = (e^{tX} \geq e^{t(1+\delta)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}} \leq \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}$$

For any $\delta > 0$, set $t = \ln(1+\delta) > 0$ to get,

$$P(X \geq (1+\delta)\mu) \leq \left\{ \frac{e^\delta}{(1+\delta)^{(1+\delta)}} \right\}^\mu$$

(ii) For any δ $0<\delta<1$, set $t = \ln(1-\delta)$ to get, <0

$$P(X \leq (1-\delta)\mu) = P(e^{tX} \geq e^{t(1-\delta)\mu}) \text{ for any } t<0$$

$$\leq \frac{E[e^{tX}]}{e^{t(1-\delta)\mu}} \leq \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu}}$$

$$\therefore P(X \leq (1-\delta)\mu) \leq \left\{\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right\}^{\mu} \quad ②.$$

Th: Let, $X_1, X_2, X_3, \ldots, X_n$ be ind. Poisson Trials, s.t., $P(X_i = 1) = p_i$ and $X = \sum_{i=1}^{n} X_i$. Let, $E[X] = \mu$. Then the following holds –

(i) for $0<\delta\leq 1$, $P(X \geq (1+\delta)\mu) \leq e^{-\mu\delta^2/3}$

(ii) for $0<\delta\leq 1$, $P(X \leq (1-\delta)\mu) \leq e^{-\mu\delta^2/2}$ ex.

(iii) for $R \geq 6\mu$, $P(X \geq R) \leq 2^{-R}$

Proof: (iii) Let, $R = (1+\delta)\mu$, Then for $R \geq 6\mu$, $\delta = \frac{R}{\mu} - 1 \geq 5$

$$P(X \geq (1+\delta)\mu) \leq \left\{\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right\}^{\mu} \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} \leq \left(\frac{e}{6}\right)^{R} \leq 2^{-R}$$

(i) We need to show, $\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \leq e^{-\delta^2/3}$ –①

$f(\delta) \to$

$$\Rightarrow \frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} - e^{-\delta^2/3} \leq 0$$

Find, $f(0), f(1), f'(\delta)\big|_{0,1}, f''(\delta)$ then thus, nature of $f(\delta)$ in $(0,1]$. (Calculus) to show $f(\delta) \leq 0$ in that range;

Take log in ①, $\delta - (1+\delta)\log(1+\delta) \leq -\delta^2/3$

$$f(\delta) = \delta^2/3 + \delta - (1+\delta)\log(1+\delta)$$

Corollary: Let, $X_1, X_2, \ldots, X_n$ be n ind. Poisson trials, st. $P(X_i=1) = p_i$ Let, $X = \sum X_i$, and $E[X] = \mu$, For $0<\delta<1$,

$$P(|X-\mu| \geq \delta\mu) \leq 2e^{-\mu\delta^2/3} \quad \text{(ex.)}$$

Ex: Let, X_i be # heads in n ind. fair coin tosses.

$$P\left(\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right)$$

Markov: $2/3$

Chebyshev: $4/n$

here, $\mu = \frac{n}{2}$, $\delta = \frac{1}{2}$

$$\therefore P\left(\left|X-\frac{n}{2}\right| \geq \frac{n}{4}\right) \leq 2e^{-n/24}$$

(inverse exponential)

$$P\left(\left|X-\frac{n}{2}\right| \geq \frac{1}{2}\sqrt{6n\ln\left(\frac{n}{2}\right)}\right) \leq \frac{4}{n}$$

$\mu\delta \to O(\sqrt{n\ln n})$ $= \frac{1}{2}\sqrt{6n\ln(n/2)}$

$e^{-n/2 \cdot \delta^2/3} = \frac{2}{n}$

$\Rightarrow \frac{n}{2}\frac{\delta^2}{3} = \ln\left(\frac{n}{2}\right)$

$\Rightarrow \delta = \sqrt{6\ln(n/2)/n}$

CLASS - 15 Dt 29/9/23

LOAD BALANCING:-

Processors : n (Identical) Jobs : m (Identical) $\lceil \frac{m}{n} \rceil$ $m = n$

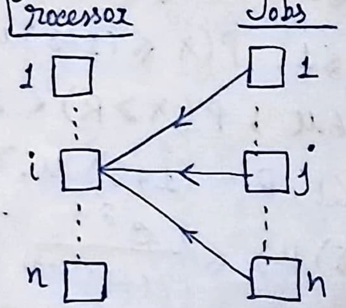
X_i : random variable denoting # jobs assigned to processor i.

$P(\max_i X_i > c) \leq ?$ ← maximum loaded processor.

$P(\exists$ a machine i, such that, $X_i > c)$

$P(X_i > c) \leq ?$

$$Y_{ij} = \begin{cases} 1, & \text{if the } j^{th} \text{ job is assigned to machine } i. \\ 0, & o/w \end{cases}$$



$X_i = \sum_j Y_{ij}$ claim:- Y_{ij}'s are ind.

Setting:- n identical jobs to be ~~arrangem~~ "assigned" to n identical machines. Jobs arrive in a stream. Assign s.t. the load is evenly balanced.

Approach:- Simply assign each jobs to one of the processes uniformly at random.

Analysis: X_i : r.v. denoting #jobs assigned to machine i.

Y_{ij} : Random Variable denoting the relation between job j and m/c i.

$$Y_{ij} = \begin{cases} 1, & \text{if job } j \text{ is assigned to m/c } i. \\ 0, & o/w \end{cases}$$; Y_{ij}'s are independent.

$P(Y_{ij} = 1) = \frac{1}{n}$

$X_i = \sum_{j=1}^{n} Y_{ij}$ $E[X_i] = n \cdot \frac{1}{n} = 1.$

We are interested in how much loaded each m/c will be.

$P(X_i > c) \leq \frac{e^{c-1}}{c^c}$ | $0 < \delta \leq 1$, $P(X \geq (1+\delta)\mu) \leq \frac{e^{\delta}}{(1+\delta)^{1+\delta}}$ From chernoff bound,

In order for there to be a small probability of any X_i exceeding c, we will take the union bound over ~~i~~ $i = 1, 2, \ldots, n$ and so we need to choose c large enough to drive $P(X_i > c)$ down well below $1/n$ for -

-each i.

Goal:- Make c^c large enough.

To understand the growth of c^c, we try to find x, s.t. $x^x = n$. Let $f(n)$ denote x. ——— ①

$$x \log x = \log n \quad \text{——— ②}$$

$$\Rightarrow \log x + \log\log x = \log\log n \quad \text{——— ③}$$

(By ③)

$$2\log x = \log x + \log x > \log x + \log\log x = \log\log n \quad \text{——— ④}$$

② ÷ ④, $\frac{x}{2} \leq \frac{\log n}{\log\log n}$ $\Rightarrow$ $\cancel{\frac{\log n}{\log\log n}}$ ——— ⑤

Also, $\frac{\log n}{\log\log n} \leq x$

$$\frac{1}{2}x \leq \frac{\log n}{\log\log n} \leq x = f(n)$$

$[x \log\log n = x(\log x + \log\log x)$
$= x\log x + x\log\log x$
$\geq x\log x = \log n]$

$$\Rightarrow f(n) = \theta\left(\frac{\log n}{\log\log n}\right)$$

$$f(n) = \theta\left(\frac{\log n}{\log\log n}\right)$$

Now, if we set $c = e \cdot f(n)$ and then apply the upper tail chernoff bound.

$$P(X_i > c) < \frac{e^{c-1}}{c^c} < \left(\frac{e}{c}\right)^c = \left(\frac{1}{f(n)}\right)^{ef(n)} \leq \frac{1}{n^2}.$$

Then, applying the union bound over the upper tail bound for $X_1, X_2, \cdots, X_n$ we have:

Th: With prob. at least $\left(1-\frac{1}{n}\right)$, no processor receives than $e \cdot f(n) = \theta\left(\frac{\log n}{\log\log n}\right)$ jobs;

modified:- $1-\frac{1}{n^c}$;

Ex: Let, $X_1, X_2, X_3, \ldots, X_n$ be ind. r.v.s with $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$;

Let $X = \sum_i X_i$. For any $\delta > 0$;

$$P(|X| \geq \delta) \leq 2e^{-\delta^2/2n}.$$

(using it) problem- Given an $n \times m$ matrix $\underset{\sim}{A} = [a_{ij}]$ with $a_{ij} \in \{0,1\}$. Let $\underset{\sim}{A} \cdot \underset{\sim}{b} = \underset{\sim}{c}$. $\underset{\sim}{b} = (b_i)$ where, $b_i \in \{-1,1\}$.

Suppose, we are looking for a matrix $\underset{\sim}{b}$ that ~~maxi~~ minimizes,

$$\|\underset{\sim}{A}\,\underset{\sim}{b}\|_\infty = \max_{i=1}^{n} |c_i|$$

Randomly and independently choose b_i's such that $P(b_i = 1) = P(b_i = -1) = \frac{1}{2}$. Then,

$$P\left(\|\underset{\sim}{A}\underset{\sim}{b}\|_\infty \geq \sqrt{4m \ln n}\right) \leq \frac{2}{n}.$$

Dt: 03/10/23

CLASS - 15

Application of Chernoff :-

Problem Estimate the probability that a particular gene mutates in a given population. Given a DNA sample, a lab test can determine if it carries the mutation. As the test is expensive, we want a relatively reliable estimate from a small # of samples.

Soln. Let, p : unknown value we are trying to estimate. Let us have n samples and $X (= n\hat{p})$ of these n samples have the mutation.

Definition (Confidence interval): A $1-\alpha$ confidence interval for a parameter p is an interval $[\hat{p}-\delta, \hat{p}+\delta]$ such that, $P(p \in [\hat{p}-\delta, \hat{p}+\delta]) \geq 1-\alpha$. We want both the interval 2δ and the error probability α is to be small, so, we derive a trade off with n.

X follows $\sim Bin(n, p)$; $E[X] = np$.

$$P(p \in [\hat{p}-\delta, \hat{p}+\delta]) = P(np \in [n(\hat{p}-\delta), n(\hat{p}+\delta)]) \geq 1-\alpha$$

If $p \notin [\hat{p}-\delta, \hat{p}+\delta]$, we have the following two events :-

(i) if $p < \hat{p}-\delta$, then $X = n\hat{p} > n(p+\delta) = E[X] \cdot \left(1+\frac{\delta}{p}\right)$;

(ii) if $p > \hat{p}+\delta$, then $X = n\hat{p} < n(p-\delta) = E[X] \cdot \left(1-\frac{\delta}{p}\right)$;

Apply chernoff bounds,

$$P(p \notin [\hat{p}-\delta, \hat{p}+\delta]) = P\left(X < np\left(1-\frac{\delta}{p}\right)\right) + P\left(X > np\left(1+\frac{\delta}{p}\right)\right)$$

$$\leq e^{-np\left(\frac{\delta}{p}\right)^2/2} + e^{-np\left(\frac{\delta}{p}\right)^2/3}$$

$$\leq e^{-n\delta^2/2p} + e^{-n\delta^2/3p}$$

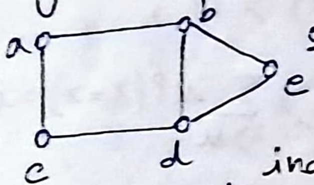
Set this to α; $\leq e^{-n\delta^2/2} + e^{-n\delta^2/3}$ [as $p \leq 1$].

PROBABILISTIC METHOD

Disjoint Union: $\biguplus$

Max Cut :- $G = (V, E)$, we want a partition $V = A \uplus B$ s.t. the no. of edges going across the cut (A, B) is maximized. $(A \cap B = \phi)$ (As of now no polynomial order algorithm is there to find max cut of a graph);

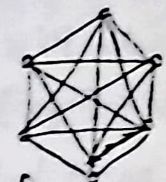
Independent Set :- $G = (V, E)$, $V' \subseteq V$ such that there is no edges $e \in E$, between any pair of vertices belonging to V';



e.g.: (a, d), (a, e), (b, c), (c, e)

For, finding max independent set is also ~~Kn~~ a ~~[illegible]~~ np-complete problem with no polynomial time algorithm;

K_n



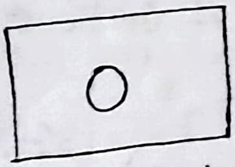
Size of independent set = 1.

K_n color edges and we are interested in "Sizes" of monochromatic subclique.

Th: If $\binom{n}{k} 2^{-\binom{k}{2}+1} < 1$, then it is possible to color the edges of K_n with two colours so that it has no monochromatic K_k subgraph.

Pf: Colour the edges uniformly at random.

Ω:



$P(\cdot) > 0$

Let, E_i be the event that clique i of size k is monochromatic.

$i = 1, 2, 3, \ldots, \binom{n}{k}$.

We are interested in a structure of certain property.

* Probabilistic methods are existential proof.

Probability of a k size ~~subgraph~~ clique is monochromatic (red/blue) $= 2 \times \left(\frac{1}{2}\right)^{\binom{k}{2}} = P(E_i)$

There are $\binom{n}{k}$ many cliques.

We have to find $P\left(\bigcap_{i=1}^{\binom{n}{k}} \overline{E_i}\right) = P\left(\overline{\bigcup_{i=1}^{\binom{n}{k}} E_i}\right)$

$= 1 - P\left(\bigcup_{i=1}^{\binom{n}{k}} E_i\right)$

$\geqslant 1 - \sum_{i=1}^{\binom{n}{k}} P(E_i)$ (By union bound)

$= 1 - \binom{n}{k} 2 \times \left(\frac{1}{2}\right)^{\binom{k}{2}}$

$= 1 - \binom{n}{k} 2^{-\binom{k}{2}+1}$ $\left(\because \binom{n}{k} 2^{-\binom{k}{2}+1} < 1 \Rightarrow 1 - \binom{n}{k} 2^{-\binom{k}{2}+1} > 0\right)$

$\therefore P\left(\bigcap_{i=1}^{\binom{n}{k}} \overline{E_i}\right) > 0$

EXPECTATION ARGUMENT :- Let X be r.v. and $E[X]=\mu$ then $P(X \geq \mu) > 0$ and $P(X \leq \mu) > 0$

(By Contradiction). Let $P(X \geq \mu) \not> 0 \Rightarrow P(X \geq \mu) = 0$

Then, $\mu = E[X] = \sum_x x P(X=x) = \sum_{x<\mu} x P(X=x) < \sum_{x<\mu} \mu \cdot P(X=x) = \mu$

which can't hold, So, $P(X \geq \mu) \neq 0 \Rightarrow P(X \geq \mu) > 0$;

Again let, $P(X \leq \mu) = 0$

Then, $\mu = E[x] = \sum_x x P(X=x) = \sum_{x>\mu} x P(X=x) > \sum_{x>\mu} \mu P(X=x) = \mu$

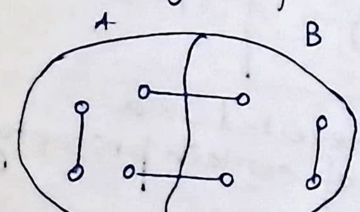
which can't hold, So, $P(X \leq \mu) > 0$;

FINDING LARGE CUTS :-

Th: Let $G=(V,E)$ be an undirected graph with, $|V|=n$, and $|E|=m$. There exists a cut of size at least ($\geq$) $m/2$?

Proof: Generate a random partition. Pick a vertex and randomly and independently put it to either A or B. Let the edges of G be $e_1, e_2, \ldots, e_i, \ldots, e_m$.

$= (v_x, v_y)$



Let, $X_i = \begin{cases} 1, & \text{if edge } e_i \text{ goes across the partition} \\ 0, & \text{o/w.} \end{cases}$ $\quad i = 1(1)m$

$P(X_i = 1) = P$ (edge e_i connects a vertex in A to a vertex in B)

$= 1/2$

$E[X_i] = \frac{1}{2}$

$C(A,B)$: A r.v. denoting the size of the cut induced by the partition A, B.

$C(A,B) = \sum_{i=1}^{m} X_i \qquad E[C(A,B)] = \frac{m}{2}$

By the previous theorem, $\exists$ a cut of size at least $m/2$;

Ex: Show that the expected # sample drawn* before finding a cut with value of at least $m/2$ is $(m/2 + 1)$

*: No. of runs of the experiment

Soln: Let p: Success probability i.e., $p = P\left(c(A,B) \geq m/2\right)$ and $c(A,B) \leq m$.

$$\frac{m}{2} = E\left[c(A,B)\right] = \sum_{i<\frac{m}{2}} i\, P\left(c(A,B)=i\right) + \sum_{i\geq \frac{m}{2}} i\cdot P\left(c(A,B)=i\right)$$

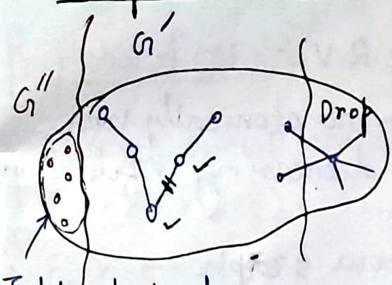
$$\leq \left(\frac{m}{2}-1\right)(1-p) + mp$$

$$\Rightarrow \cancel{\frac{m}{2}} \leq \cancel{\frac{m}{2}} - \frac{mp}{2} - 1 + p + mp$$

$$\Rightarrow \frac{m}{2}p - 1 + p \geq 0 \rightarrow p\left(\frac{m}{2}+1\right) \geq 1 \Rightarrow p \geq \frac{1}{\left(\frac{m}{2}+1\right)}$$

$\therefore 1/p \leq \frac{m}{2}+1$ ✓ (waiting for success)

Independent Set:-



Independent set.

$G \xrightarrow[\text{vertices}]{\text{drop}} G'$

↑ is a random process;

$G' \rightarrow G''$ (deterministic process)

drop an edge and any one vertex corresponding to it.

Class – 167 Dt. 6/10/23

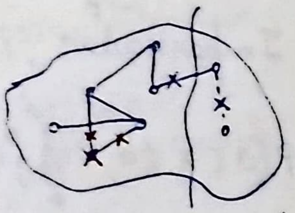
12/10 (THU) : 11:20 a.m. Tutorial

18/10 (WED) : 6:30 pm. Test.

Sample and Modify:-

$|V| = n$, $|E| = m$;

Th: Let $G=(V,E)$ be a connected graph on n vertices with $m \geq \frac{n}{2}$ edges. Then, G has an independent set with at least $n^2/4m$ vertices; (Graph is suitably dense)



Pf: Step-1:- Sample:- Retain vertices in G independently with probability α; i.e., delete vertices (and edges incident on it) in G with prob. $1-\alpha$.

Step-2:- Modify:- Drop the remaining edges, and for each edge dropped drop one of its endpoint;

Let, X: random variable denoting the vertices that were retained in step 1;

Y: random variable that denotes the # edges that survived after ~~ste~~ step 1;

$$E[X] = n\alpha$$

An edge survived after step 1 if and only if both of its endpoints survive. $E[Y] = m\cdot\alpha^2$

Generating random numbers has significant cost;

Difference b/w arbitrary & random;

The 2nd step deletes all edges and at most Y vertices.

We are sure that the vertices retained form an independent set S ;

$E[\#\text{vertices in } S] \geq E[X-Y] = E[X]-E[Y] = n\alpha - m\alpha^2 = f(\alpha)$

$f'(\alpha) = n - 2m\alpha = 0 \Rightarrow \alpha = \frac{n}{2m}$ as $m \geq \frac{n}{2} \Rightarrow \alpha \leq 1$

$f''(\alpha) = -2m < 0$

$\therefore E[\#\text{ vertices in } S] \geq \frac{n^2}{2m} - \frac{n^2}{4m} = \frac{n^2}{4m}$ (completes argument)

$\frac{1}{n} \times \sum \deg(v) = 2m \times \frac{1}{n} \Rightarrow d_{avg} = \frac{2m}{n}$ average degree!

$\therefore \alpha = \frac{1}{d_{avg}}$ ✓

CONTINUOUS RANDOM VARIABLE :-

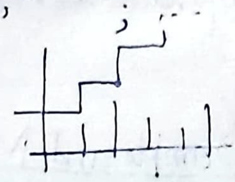
Cumulative Distribution Function for Discrete R.V. :-

$CDF(K) = P(X \leq K) = \sum_{-\infty}^{K} pmf_X(x)$ ↳ monotonically non decreasing function;

$P(X=K) = CDF(K) - CDF(K-1)$

* Graph of CDF of DRV → Piecewise continuous graph ;

CRV → Smooth Continuous " ;



⇒ A r.v. X is said to be continuous if there is a non-ve function f_X called the prob. density function (pdf) of X s.t.,

$P(X \in B) = \int_B f(x)\,dx$ for every subset B of the real line.

In particular, $P(a \leq X \leq b) = \int_a^b f_X(x)\,dx$ [area under the curve].

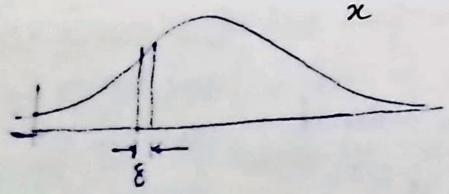
$\Rightarrow P(a \leq X \leq a) = \int_a^a f(x)\,dx = 0$

$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$. For $f_X(x)$ to be a pdf, it has to satisfy the following conditions —

① $f_X(x) \geq 0 \;\forall x$ and $\int_{-\infty}^{\infty} f_X(x)\,dx = P(-\infty < X < \infty) = 1$

$P(X \in [x, x+\delta]) = \int_x^{x+\delta} f_X(t)\,dt \approx f_X(x) \cdot \delta$.

↳ as the prob. mass / unit length near x.



[6] (PDF : can take arbitrarily large values)

Consider a CRV, X with PDF

$$f_X(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & \text{if } 0 < x \leq 1 \\ 0, & \text{o/w} \end{cases}$$

figure out what is: $\int_{-\infty}^{\infty} f_X(x)\,dx = ?$

$\rightarrow \int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 \frac{1}{2\sqrt{x}}\,dx = [\sqrt{x}]_0^1$

$= 1$ (qualifies as PDF but takes arbitrarily large values as x is near 0)

Cumulative Distribution Fn:-

$$F_X(x) = P(X \leq x) = \begin{cases} \sum_{i \leq x} pmf_X(i), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{x} f_X(t)\,dt, & \text{if } X \text{ is continuous.} \end{cases}$$

Properties of CDF:-

- $F_X(x)$ is monotonically non decreasing if $x \leq y$, then $F_X(x) \leq F_X(y)$;
- $F_X(x) \to 0$ as $x \to -\infty$ and $F_X(x) \to 1$ as $x \to \infty$
- If X is discrete, then $F_X(x)$ is a piecewise constant function of x;
- If X is continuous, then $F_X(x)$ is a continuous function of x;
- If X is discrete, $pmf_X(x) = F_X(K) - F_X(K-1)$
- If X is continuous, then PDF and CDF can be obtained from each other by integration or differentiation;

$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \qquad f_X(x) = \frac{d}{dx} F_X(x)$ ← (should exist)

For a continuous R.V.,

$$P(a - \varepsilon/2 \leq X \leq a + \varepsilon/2) = \int_{a-\varepsilon/2}^{a+\varepsilon/2} f_X(x)\,dx \approx \varepsilon \cdot f(a).$$

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

You can define the n^{th} moment also, $E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\,dx.$

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

- We will say the expectation is well defined if,

$\int_{-\infty}^{\infty} |x| f_X(x)\,dx < \infty$; (i.e., converges)

Ex: Consider a CRV with pdf $f_X(x) = \frac{c}{(1+x)^2}$ where, c is a constant such that $f_X(x)$ qualifies to be a pdf ;

Look at $\int_{\infty}^{\infty} |x| \frac{c}{1+x^2} dx \leq \int_{-\infty}^{\infty} |x| \frac{c}{x^2} dx = \int_{-\infty}^{\infty} \frac{c}{|x|} dx$

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

Now, we know, $\int_{1}^{\infty} \frac{dx}{x}$ diverges.

$= -\cancel{\frac{c}{2} \log(1+x^2)}$

∴ Given function ~~has~~ doesn't have well defined expectation.

Ex: Let, $X_1, X_2, \ldots, X_n$ be r.v.s (continuous/discrete). If the events $\{X_1 \leq x\}, \ldots, \{X_n \leq x\}$ are independent for every x, then the CDF of $X = \max\{X_1, X_2, \ldots, X_n\} = ?$

→ $F_X(x) = F_{X_1}(x) \cdot F_{X_2}(x) \cdots F_{X_n}(x)$ (To prove)

Pf: $F_X(x) = P\left((X_1 \leq x) \cap (X_2 \leq x) \cap (X_3 \leq x) \cap \ldots \cap (X_n \leq x)\right)$

$P(X \leq x)$

$= P(X_1 \leq x) \cdot P(X_2 \leq x) \cdots P(X_n \leq x)$

$= F_{X_1}(x) \cdot F_{X_2}(x) \cdots F_{X_n}(x)$

Note:- Nice way to ~~get~~ get pmf/pdf from ~~CDF~~ this ;

Exponential Random Variable :- An exponential random variable has a pdf of form -

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0 \\ 0, & \text{o/w} \end{cases}$$

Ex: (i) verify $f_X(x)$ is a pdf ; (ii) $E[X] = ?\ \frac{1}{\lambda}$ (iii) $Var[X] = ?\ \frac{1}{\lambda^2}$

Uniform Random Variable :- $X \sim U[a, b]$

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b \\ 0, & \text{o/w} \end{cases}$$

Ex: (i) Verify $f_X(x)$ is a pdf ; (ii) $E[X] = ?\ \left(\frac{a+b}{2}\right)$ (iii) $Var[X] = ?\ \frac{(b-a)^2}{12}$

Geometric R.V. :-

$X \sim Geo(p)$ $P(X=k) = p(1-p)^{k-1}$; CDF of X: $P(X \leq n) = ?$

$F_{exp}(x) = P(X \leq 0) = 0$ for $x \leq 0$;

For $x>0$, $F_{exp}(x) = \int_0^x \lambda e^{-\lambda t}\, dt$

$= \dfrac{\lambda e^{-\lambda t}}{-\lambda}\Big|_0^x = 1-e^{-\lambda x}$

$F_{Geo}(n) =$

$F_{Geo}(n) = \underline{1-(1-p)^n}$

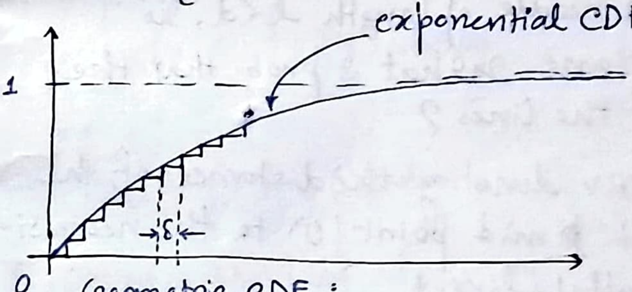
$n = 1, 2, 3, \ldots$

Now,

$\Rightarrow 1-(1-p)^n = 1-e^{-\lambda x}$

$\Rightarrow n \ln(1-p) = -\lambda x$

$\Rightarrow -\dfrac{\ln(1-p)}{\lambda} = \dfrac{x}{n} = \delta$

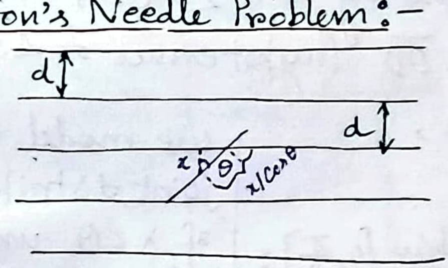


Geometric CDF :
$1-(1-p)^n$ with $p = 1-e^{-\lambda\delta}$

* In continuous R.V., relate via CDF not PMF;

* thats why we use continuous sometimes in study of discrete to use power of calculus;

Buffon's Needle Problem:-



▸ Stick/needle of length l randomly place that on the table. What is the prob. that the stick intersect the lines ? $l<d$

↳ Concept of joint dist. of CRV reqd. ;

JOINT PDFs OF MULTIPLE RVs:- X, Y are C.R.V ;

PDF: $f_{X,Y}(x,y)$: non negative function ;

$$P((X,Y)\in B) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\, dy\, dx$$

For every subset B of $\mathbb{R}^2$;

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

$P(a \leq X \leq a+\varepsilon,\ c \leq Y \leq c+\varepsilon)$

$$= \int_{c}^{c+\varepsilon}\int_{a}^{a+\varepsilon} f_X(x,y)\,dx\,dy \approx f_{X,Y}(x,y)\,\varepsilon^2$$

← Averaged area at (x,y);

<u>Marginal:-</u>

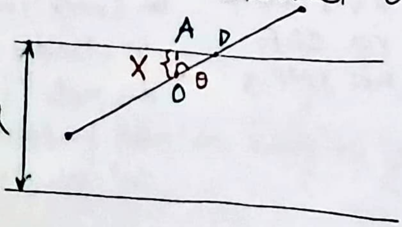
$$P(X \in A) = P(X \in A \text{ and } Y \in \{-\infty, +\infty\})$$

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

Compare with, $\int_A f_X(x)\,dx = P(X \in A)$, and observe that the marginal pdf f_X of X is given by $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$,

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

<u>Buffon's Needle Problem</u>:- A 2D-plane is ruled by parallel lines 'd' distance apart. A needle of length $l<d$, is randomly thrown on the plane. What is prob. that the needle will intersect one of the lines?



X: CRV denoting the distance of the stick ~~p~~ mid point (O) to the nearest parallel line;

θ: CRV denoting the $\angle AOD$;

The needle will intersect a line iff $\overline{OD}$ (hypotenuse of $\triangle OAD$) is less than $l/2$ i.e., $\frac{X}{\cos\theta} \leq l/2$ i.e.,

$X \leq l/2 \cos\theta$.

X varies between $[0, \frac{d}{2}]$, θ varies b/w $[0, \frac{\pi}{2}]$; | we model the joint distribution of X & θ uniform in ~~range~~ their respective range;

Fix, $f_{X,\theta}(x,\theta) = \begin{cases} \frac{4}{\pi d}, & (x,\theta) \in [0, \frac{d}{2}] \times [0, \frac{\pi}{2}] \\ 0, & o/w. \end{cases}$

$$P(\text{intersection}) = P(X \leq \tfrac{l}{2}\cos\theta)$$

$$= \int_0^{\pi/2}\int_0^{l/2\cos\theta} \frac{4}{\pi d}\,dx\,d\theta$$

$=$

JOINT CDFs:- $F_X(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\,dt\,ds$ ~~dy dx~~

X,Y Conversely, the PDF can be recovered from the CDF as, $f_{X,Y}(x,y) = \frac{\delta^2}{\delta x \delta y}\left(F_{X,Y}(x,y)\right)$

▷ CRV $\begin{cases} X \sim U[0,1] \\ Y \sim U[0,1] \end{cases}$ $X+Y \sim$

X and Y are independent

$F_{X+Y}(a) = P(X+Y \le a) = \iint_{x+y \le a} f_X(x) f_Y(y)\,dx\,dy$.

$= \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(x) f_Y(y)\,dx\,dy = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{a-y} f_X(x)\,dx\right) f_Y(y)\,dy$

$= \int_{-\infty}^{\infty} F_X(a-y)\cdot F_Y(y)\,dy$

To Obtain the pdf f_{X+Y} of $X+Y$, we do

$f_{X+Y}(a) = \frac{d}{da}\left[\int_{-\infty}^{\infty} F_X(a-y)\cdot f_Y(y)\,dy\right]$

$= \int_{-\infty}^{\infty} \frac{d}{da} F_X(a-y) f_Y(y)\,dy = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y)\,dy$

Using the expression for $f_X(a)$ and $f_Y(a)$.

$f_{X+Y}(a) = \int_0^1 f_X(a-y)\,dy$.

$\left.\begin{array}{l} \text{For, } 0 \le a \le 1,\ f_{X+Y}(a) = \int_0^a dy = a \\ \text{For, } 1 \le a \le 2,\ f_{X+Y}(a) = \int_{a-1}^{1} dy = 2-a \end{array}\right\}$ ex.

→ $f_X(a) = f_Y(a) = \begin{cases} 1, & 0 \le a < 1 \\ 0, & \text{o/w.} \end{cases}$

$\left.\begin{array}{l} 0 \le a \le 1 \\ 0 \le y \le 1 \end{array}\right\} -1 \le a-y \le a \le 1$

$-1 \to 0$: No mass!

$\left.\begin{array}{l} 1 < a \le 2 \\ 0 \le y \le 1 \end{array}\right\}$

$a-1 \le a-y \le 1 \le 2$

$$f_{X+Y}(a) = \begin{cases} a & , 0 \leq a < 1 \\ 2-a & , 1 \leq a \leq 2 \\ 0 & , o/w \end{cases}$$

DERIVED DISTRIBUTION :- Let, X be a CRV and let $Y = g(X)$. Given the PDF of X, how to calculate the PDF of Y?

Step ① :- Calculate the CDF of Y :-

$$F_Y(y) = P(g(X) \leq y) = \int_{\{x \mid g(x) \leq y\}} f_X(x)\, dx$$

② Differentiate wrt y to obtain the pdf of Y :-

$$f_Y(y) = \frac{d}{dy}(F_Y(y)) = \frac{d}{dy}\left(\int_{\{x \mid g(x) \leq y\}} f_X(x)\, dx \right)$$

Ex: Let $X \sim U[0,1]$ and $Y = \sqrt{X}$

$$F_Y(y) = P(\sqrt{X} < y) = \int_0^1$$

Ex:- Let X be a CRV, $Y = aX + b$, Y is linear fn. of R.V. of X. Compute PDF of Y in terms of pdf of X; $(a \neq 0)$; A $f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

Ex:- Let X follows exp. R.V. & $Y = aX + b$. Figure out PDF of Y. Check whether it'll be always exp. R.V. ;

Problem sheet 3

Dt: 12/10/23

Thm: If X & Y are indep. random var., $MGF_{X+Y}(t) = MGF_X(t) \cdot MGF_Y(t)$:-

$$E[e^{t(x+y)}] = E[e^{tx} \cdot e^{ty}] = E[e^{tx}] \cdot E[e^{ty}]$$
$$= MGF_X(t) \cdot MGF_Y(t)$$

By induction extend it ;

Thm:- Sum of independent poisson trials, $X = \sum_{i=1}^{n} X_i$

$$MGF_X(t) = \prod_{i=1}^{n} MGF_{X_i}(t) = e^{(e^t-1)\underbrace{\sum_{i=1}^{n} p_i}_{\mu}}$$

$[1+x \leq e^x]$

$$\rightarrow E[e^{tx}] = pe^t + (1-p) = 1 + (e^t - 1)p \leq e^{(e^t-1)p}$$

* we have chosen poisson to get good exp. for MGF ;

$$E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$$
$$= \sum_{i=1}^{n} p_i$$

2] $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$

$$MGF_{X_i}(t) = \frac{e^t}{2} + \frac{e^{-t}}{2}$$

$$\rightarrow e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots$$

$$e^{-t} = 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \dots$$

$$e^t + e^{-t} = 2\left(1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \dots\right)$$

$$MGF_X(t) = e^{t^2 n/2}$$
$$\Uparrow$$

$$\Rightarrow \frac{1}{2}(e^t + e^{-t}) = \sum_{i \geq 0} \frac{t^{2i}}{(2i)!} \leq \sum_{i \geq 0} \frac{\left(\frac{t^2}{2}\right)^i}{i!} = e^{t^2/2}$$

* In Q1 → all are multiplicative chernoff bound. → $(1+\delta)\mu$
But in Q2 → $\mu = 0$ ∴ mult. cher. bound don't work here, we use additive chernoff bound. → $X \geq a$..

$P(X \geq a)$ ← constant :- $\mu + K$

$$\leq \frac{E[e^{tx}]}{e^{ta}}, t > 0$$

$$\leq e^{\frac{t^2 n}{2} - ta} \qquad t = \frac{a}{n}$$

$$= e^{\frac{a^2}{2n} - \frac{a^2}{n}}$$

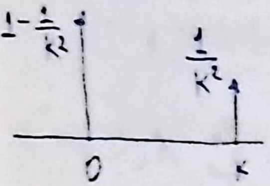
$$= e^{-\frac{a^2}{2n}}$$

* mult chernoff bound → sum of ux poisson trials
additive " " → used when $\mu = 0$

] An inequality is tight when it is equality in some cases;

Markov- $P(X \geq a) = \frac{E[X]}{a}$

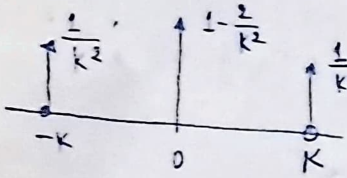
$$P(X) = 1 - \frac{1}{K^2}, \ X = 0$$
$$= \frac{1}{K^2}, \ X = K$$
$$= 0, \ \text{o.w.}$$



$$E[X] = \frac{1}{K}$$

$$P(X \geq K) = \frac{1}{K^2} = \frac{\frac{1}{K}}{K} = \frac{E[X]}{K}$$

* Chebyshev is tighter when variation in both side (i.e, not reqd to be non-ve) :-



$$E[X] = 0$$
$$Var[X] = 2$$
$$P(|X| \geq K) = \frac{2}{K^2} = \frac{Var[X]}{K^2}$$

Chebyshev is tight;

17] $Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] - \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)$

We'll show $Cov(X_i, X_j) = 0$ & prove!

18] Chebyshev w/o modulus!

$$P(\underbrace{X - E[X]}_{Y} \geq \underbrace{t\sigma(X)}_{a}) \leq \frac{1}{1+t^2}$$

$E[Y] = 0$

$\sigma_Y = \sigma_X = \sigma$ (say)

$P(Y \geq a)$ ← non-ve R.V.

$= P((Y+K)^2 \geq (a+K)^2)$ markov.

$\leq \frac{E[(Y+K)^2]}{(a+K)^2}$ → $E[Y^2 + 2YK + K^2]$

$= E[Y^2] + 2K\,E[Y] + K^2$

$= \sigma^2 + K^2$

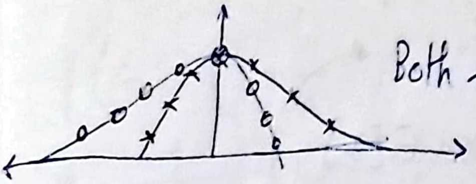
$= \frac{\sigma^2 + K^2}{(a+K)^2}$

differentiate to minimize & get, $K = \frac{\sigma^2}{a}$

$$= \frac{\sigma^2 + \frac{\sigma^4}{a^2}}{\left(a + \frac{\sigma^2}{a}\right)^2} = \frac{\frac{\sigma^2}{a}\left(a + \frac{\sigma^2}{a}\right)}{\left(a + \frac{\sigma^2}{a}\right)^2} = \frac{\sigma^2}{a^2 + \sigma^2}$$

$\Rightarrow P(X - E[X] \geq a) \leq \dfrac{\sigma^2}{a^2+\sigma^2}$ ✓ cantelli's ineq.

$P(X - E[X] \leq -a) \leq \dfrac{\sigma^2}{a^2+\sigma^2}$



Both have same chebyshev bound! but diff. bound in two sides we can't break down chebyshev;

From chebyshev,

$P(|X-E[X]| \geq a) \leq \dfrac{\sigma^2}{a^2}$

$P(|X-E[X]| \geq t\sigma) \leq \dfrac{1}{t^2}$ — (i)

Adding two sides of prev:-

$P(|X-E[X]| \geq t\sigma) \leq \dfrac{2}{1+t^2}$ — (ii)

when (ii) is better than (i),

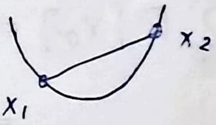
$\dfrac{2}{1+t^2} < \dfrac{1}{t^2} \Rightarrow 2t^2 < 1+t^2 \Rightarrow t^2 < 1 \Rightarrow t < 1$.

But $\dfrac{1}{t^2}$ & $\dfrac{2}{1+t^2}$ are both > 1, that is trivial bound for Probabilities. Thus, for both sided case, it doesn't help at all;

18. $Pr[X \neq 0] = Pr[X \geq 1] \leq \dfrac{E[X]}{1^2} = E[X]$ As X is non negative integer valued;

→ When f is convex function,

$E[f(X)] \geq f(E[X])$



→ $E[X \mid X \neq 0] = \sum_{X \geq 1} X \cdot \dfrac{P(X)}{P(X \neq 0)}$

$\dfrac{E[X]^2}{E[X^2]} \leq P(X \neq 0)$.

$= \dfrac{\sum_{X \geq 1} X \cdot P(X)}{P(X \neq 0)} = \dfrac{E[X]}{P(X \neq 0)}$

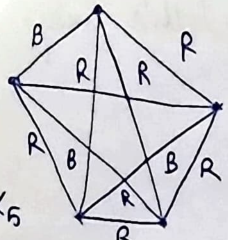
$\Rightarrow E[X^2 \mid X \neq 0] = \dfrac{E[X^2]}{P(X \neq 0)}$

Now, $E[X^2] \geq E[X]^2 \Rightarrow E[X^2 \mid X \neq 0] \geq E[X \mid X \neq 0]^2$

$\Rightarrow \dfrac{E[X^2]}{P(X \neq 0)} \geq \dfrac{E[X]^2}{P(X \neq 0)^2} \Rightarrow \dfrac{E[X]^2}{E[X^2]} \leq P(X \neq 0)$;

12. Generalised qn of class problem;

$P(\ ) > 0$
$E[X] > K$
$\exists\ X > E[X]$



K_5

$\vdots$

K_n

$\exists$ at most ${}^nC_a \cdot 2^{1-{}^aC_2}$ many K_a (monochromatic).

$X = \#$ monochromatic K_a

$E[X] = {}^nC_a\, 2^{1-{}^aC_2}$

$\exists$ a coloring for which, $X < E[X]$

$X = \#$ monochromatic K_a

$\# K_a = {}^nC_a \qquad Pr = \left(\frac{1}{2}\right)^{{}^aC_2} \cdot 2$

$\therefore E[X] = {}^nC_a \cdot 2^{1-{}^aC_2}$;

10 Tournament $\to$ directed complete graph

$\circ \to \circ \to \circ \to \circ \to \circ$ (Hamiltonian path)

* covers all vertices once

<u>Lemma:</u> Any tournament has a hamiltonian path ;

(en) (Induction)

$X = \#$ Hamiltonian paths. No. of tournaments $= 2^{\binom{n}{2}}$

$n! \to$ no. of permutations ; $\hookrightarrow$ a tournament is selected randomly ;

$P($a <u>particular</u> permutation $\sigma \in T) = \left(\frac{1}{2}\right)^{n-1}$

$\hookrightarrow$ <u>hamiltonian path</u>

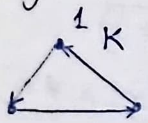
$\circ \to \circ \to \circ \to \circ \cdots \to \circ$

$(n-1)$ edges prob. of each in this particular direction is $\frac{1}{2}$;

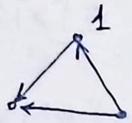
$E[X] = \sum_\sigma E[X_\sigma] = \sum_\sigma P(\sigma \in T)$

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

9 S_K *: For every set of size K $\exists$ a player who defeats all K of them.



S_1 ✓ $T_3 \not\models S_1$



$T_3 \not\models S_1$

$v \in V \setminus K$ $T_3 \models S_1$

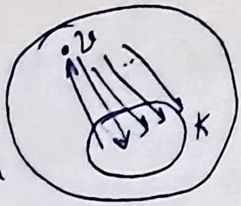
$v \to a$ s.t. $a \in K$

If ${}^nC_K (1 - 2^{-K})^{n-K} < 1$ then $\exists$ a tournament on n vertices T_n, $T_n \models S_K$.

$P(\exists T_n \mid T_n \models S_n) > 0$

$P(\nexists T_n \mid T_n \models S_n) < 1$

$\mathcal{E}_K$: for $K \subseteq V$, there is no vertices in $V \setminus K$ that defeats K



$P(\mathcal{E}_K) = \left(1 - \left(\frac{1}{2}\right)^K\right)^{n-K}$

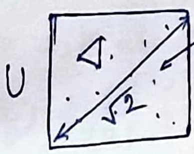
$\hookrightarrow$ v doesn't defeat all of them in K;

$P($there is no K s.t. $\exists v \in V \setminus K$ which defeats all of $K)$

$= P(\bigcup_K \mathcal{E}_K) \leqslant \sum_K P(\mathcal{E}_K) = {}^nC_K \left(1 - \frac{1}{2^K}\right)^{n-K} < 1$

Reqd result.

13



n vertices

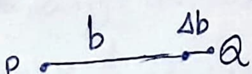
area of smallest Δ T? $\geqslant \frac{1}{100n^2}$

$\rightarrow {}^nC_3$ Δs.

Claim: $P(\Delta < \varepsilon) \leqslant 16\pi\varepsilon$

$P(\Delta < \varepsilon) \leqslant \cdots$

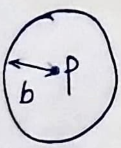
$\geqslant 0$

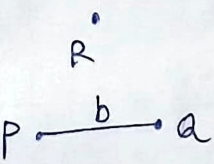


$P(PQ = b)$

$\hookrightarrow \leqslant P(b \leqslant PQ \leqslant b + \Delta b)$

$= \pi(b+\Delta b)^2 - \pi b^2 = (2\pi b)db$

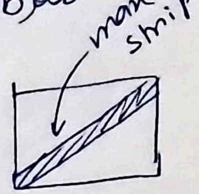


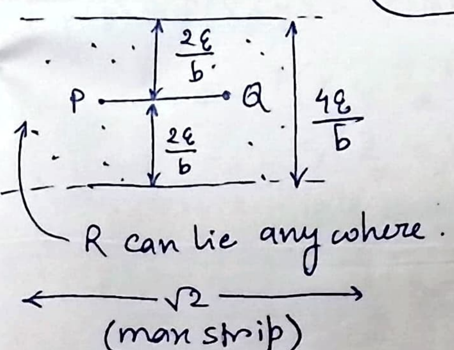


Area $< \varepsilon$

$h < \frac{2\varepsilon}{b}$

height





$$P(\Delta < \varepsilon) \leqslant \int_{b=0}^{\sqrt{2}} 2\pi b \, db \cdot \frac{4\varepsilon}{b} \times \sqrt{2}$$

$$\leqslant 4\pi\varepsilon \, [b]_0^{\sqrt{2}} \times \sqrt{2} \times 2$$

$$= 8\pi\varepsilon \times 2 = 16\pi\varepsilon$$

Class - 19 Dt 13/010/23

Ex: Let $X \sim U[0,1]$ and $Y = \sqrt{X}$. Find the pdf of Y. For every $y \in [0,1]$;

$\rightarrow F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2)$

$= \int_0^{y^2} 1 \, dx = y^2$

PDF of Y $= f_Y(y) = \frac{d}{dy} F_Y(y) = \begin{cases} 2y, & 0 \leq y \leq 1 \\ 0, & \text{o.w.} \end{cases}$

The PDF of a linear function of a R.V. :-

Let X be a c.r.v. with PDF $f_X(x)$ and let $Y = aX + b$, where a, b are scalars with $a \neq 0$. Then,

$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

Proof: Case I $(a > 0)$:-

$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$

Now, differentiate w.r.t y,

$f_Y(y) = \frac{d}{dy}(F_Y(y)) = \frac{d}{dy}\left(F_X\left(\frac{y-b}{a}\right)\right) = \frac{d}{dx}\left(F_X\left(\frac{y-b}{a}\right)\right) \cdot \left(\frac{dx}{dy}\right) \nearrow \frac{1}{a}$

$= \frac{1}{a} \frac{d}{dx}\left(F_X\left(\frac{y-b}{a}\right)\right)$

So, $f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$

Case II : $(a < 0)$:-

$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \geq \frac{y-b}{a}\right)$

$= 1 - P\left(X < \frac{y-b}{a}\right)$

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

Now, differentiate w.r.t y,

$f_Y(y) = \frac{d}{dy}\left\{1 - F_X\left(\frac{y-b}{a}\right)\right\} = -\frac{d}{dy}\left(F_X\left(\frac{y-b}{a}\right)\right) = -\frac{d}{dx}\left(F_X\left(\frac{y-b}{a}\right)\right) \frac{dx}{dy}$

$= -\frac{1}{a} \frac{d}{dx}\left(F_X\left(\frac{y-b}{a}\right)\right)$

So, $f_Y(y) = -\frac{1}{a} f_X\left(\frac{y-b}{a}\right)$

A c.r.v. X is said to follow a normal distribution with parameter μ and σ / σ^2. $(\sigma > 0)$;

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

Jacobian -

$$\iint f(x,y)\,dx\,dy \quad x = r\cos\theta,\ y = r\sin\theta$$

$$f : \mathbb{R}^2 \to [\mathbb{R}^+, (0, 2\pi)]$$

$$\iint g(r,\theta)\, dr\, d\theta$$

$$\begin{vmatrix} \frac{\delta x}{\delta r} & \frac{\delta x}{\delta \theta} \\ \frac{\delta y}{\delta r} & \frac{\delta y}{\delta \theta} \end{vmatrix} = J$$

$\ominus$ $\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} dx = 1$

$$\boxed{\begin{aligned} y &= \frac{x-\mu}{\sigma} \\ \sigma dy &= dx \\ \Rightarrow dy &= \frac{dx}{\sigma} \end{aligned}}$$

$\Rightarrow \displaystyle\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} \cdot dy = 1$

$\Rightarrow \displaystyle\frac{1}{\sqrt{2\pi}} \cdot I = 1$

$\Rightarrow I = \sqrt{2\pi}$ (need to show)

$I = \displaystyle\int_{-\infty}^{\infty} e^{-y^2/2} \cdot dy \quad \to I^2 = \left(\int_{-\infty}^{\infty} e^{-y^2/2} \cdot dy\right)\left(\int_{-\infty}^{\infty} e^{-x^2/2} \cdot dx\right)$

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

Let, $x = r\cos\theta$, $y = r\sin\theta$; $dy\,dx = r\,d\theta\,dr$ (By Jacobian)

$I^2 = \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} \cdot dy\,dx$

$= \displaystyle\int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} \cdot r\,d\theta\,dr$

Put, $z = \frac{r^2}{2}$
$dz = r\,dr$

$= 2\pi \displaystyle\int_0^{\infty} r \cdot e^{-r^2/2} dr \quad = +2\pi \int_0^{\infty} e^{-z} dz \quad = 2\pi \left[-e^{-z}\right]_0^{\infty}$

$= 2\pi$

$\Rightarrow I = \sqrt{2\pi}$ ($\square$

$\ominus$ $E[X] = \displaystyle\int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} dx \qquad x = x - \mu + \mu$

$= \displaystyle\int_{-\infty}^{\infty} \frac{x-\mu}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} dx + \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} \cdot dx$

$= J_1 + I_2$

ex: Show that J_1 is an integral of an odd function over an interval that is symmetric about

$I_2 = \mu \displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} \cdot dx$

$= \mu \cdot 1 = \mu$

0 and it must be 0 unless the +ve & −ve part both diverge to ∞ ;

put $\frac{x-\mu}{\sigma} = z$ in I_1, $I_1 = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\, e^{-z^2/2}\, dz$

$\Rightarrow dx = \sigma\, dz$

$\Rightarrow dz = \frac{dx}{\sigma}$

Now, from 0 to ∞, $I_1 = \frac{\sigma}{\sqrt{2\pi}} \int_0^{\infty} z\, e^{-z^2/2}\, dz = \frac{\sigma}{\sqrt{2\pi}} \left[-e^{-u}\right]_0^{\infty}$

$= \frac{\sigma}{\sqrt{2\pi}} \cdot 1 < \infty$

$\therefore f(z) = z \cdot e^{-z^2/2}$

$f(-z) = -z \cdot e^{-z^2/2} = -f(z)$

$\therefore f(z)$ is a odd function.

$\therefore I_1 = 0$ and $\boxed{E[X] = \mu}$ ✓

⊖ $Var[X] = E[(X-\mu)^2] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-(x-\mu)^2/2\sigma^2}\, dx$

Substitute, $y = \frac{x-\mu}{\sigma} \Rightarrow dy = \frac{dx}{x}$

$\therefore Var[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma} e^{-(x-\mu)^2/2\sigma^2}\, dx$

$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2}\, dy$ (Integration by parts)

$= \frac{2\sigma^2}{\sqrt{2\pi}} \left[y \int_0^{\infty} e^{-y^2/2} dy \Big|_0^{\infty} + 2 \int_0^{\infty} y e^{-y^2/2} dy \right] = \frac{2\sigma^2}{\sqrt{2\pi}} [0 +$

Ex: $X \sim N(\mu, \sigma^2)$; $Y = aX + b$. Then Y follows what dist.?

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

$\rightarrow E[Y] = E[aX + b] = aE[X] + b = a\mu + b$

$E[Y^2] = E[a^2X^2 + 2abX + b^2] = a^2E[X]^2 + 2ab\,E[X] + b^2$

$\therefore Var[Y] = E[Y^2] - (E[Y])^2 = a^2 E[X]^2 + \cancel{2ab E[X]} + \cancel{b^2} - a^2 E[X]^2 - \cancel{2ab E[X]} - \cancel{b^2}$

$= a^2\{E[X^2] - E[X]^2\} = a^2\sigma^2$

Ex: $Z = \frac{X-\mu}{\sigma}$ (let) $Z \sim N(0,1)$ (Standard normal)

$E[Z] = \frac{E[X] - \mu}{\sigma} = 0$ $\quad E[Z^2] = \frac{E[X^2] - 2\mu E[X] + \mu^2}{\sigma^2}$

$$E[Z^2] - E[Z]^2 = \frac{E[X^2] - 2\mu E[X] + \mu^2 - E[X]^2 + 2\mu E[X] - \mu^2}{\sigma^2}$$

$$= \frac{\sigma^2}{\sigma^2} = 1 = Var[Z]$$

CDF of Z:-

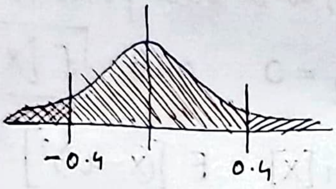
$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$

The standard normal table l values for only $\Phi(z)$ when $z \geq 0$. The others can be obtained by symmetry

$$\Phi(-0.4) = P(Z \leq 0.4) = P(Z \geq 0.4) = 1 - P(Z < 0.4)$$

$$\boxed{\Phi(-y) = 1 - \Phi(y)} \qquad = 1 - \Phi(0.4)$$

Q) $X \sim N(3, 9)$



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

$$= P\left(-\frac{1}{3} < z < \frac{2}{3}\right) = \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)$$

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

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

$$= 0.74857 + 0.62930 - 1$$

$$= 0.37787$$

Q) $P(|X-3| > 6) = ?$ $\qquad X \sim N(3, 9)$

$$= P(X - 3 < -6,\ X - 3 > 6)$$

$$= P(X < -3,\ X > 9) = P\left(Z < \frac{-3-3}{3},\ Z > \frac{9-3}{3}\right)$$

$$= P(Z < -2,\ Z > 2) =$$

Ex: X : r.v denoting the # times a fair coin flipped. Say 40 times lands heads. Say we are interested in $P(X = 20) = ?$

$$\rightarrow P(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^{40} \approx 0.1254$$

$$P(X = 20) = P(19.5 < X < 20.5) \qquad X \sim N(\mu, \sigma^2)$$

$\mu = np = 20$, $\sigma^2 = np(1-p) = 10$

$$= P\left(\frac{19.5 - 20}{\sqrt{10}} < Z < \frac{20.5 - 20}{\sqrt{10}}\right)$$

$= P(-0.16 < Z < 0.16)$

$= \Phi(0.16) - \Phi(-0.16) = \Phi(0.16) - (1 - \Phi(0.16))$

$= 2\,\Phi(0.16) - 1 = 2 \times 0.56356 - 1 = 1.12712 - 1$

$= 0.12712$ ✓

* That means normal approximates Binomial well. (~~for~~ useful for very large n);

Class - . Dt - 17/10/23

Transform of MGF :-

(n^{th} derivative of MGF at $t=0$) = n^{th} moment

→ Expectation $\xleftrightarrow{\text{Interchange}}$ Differentiation

→ MGF of a continuous RV :- $\int$ ⟷ different

Why MGF? → We can extract distribution using MGF
We can handle sum of r.v. ;

Let scalar parameter s, $M_X(s) = E[e^{sx}] = \begin{cases} \sum_x e^{sx} \cdot pmf(x) \; ; x \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{sx} \cdot f_X(x)\, dx \; ; x \text{ is continuous} \end{cases}$

Transform of MGF :-
The transform is a function of the parameter s, $M_X(s)$ is only defined for those values of s for which $E[e^{sx}]$ is finite

ex :- Transform of exponential r.v. :- $f_X(x) = \lambda e^{-\lambda x}, x \geq 0$

$M_X(s) = E[e^{sx}] = \int_0^\infty e^{sx} \cdot \lambda e^{-\lambda x}\, dx \quad , x \geq 0$

$M_X(s) = \lambda \cdot \int_0^\infty e^{(s-\lambda)x}\, dx \quad , x \geq 0$

$= \frac{\lambda}{s-\lambda} \left[e^{((s-\lambda)x)} \right]_0^\infty$ when, $s < \lambda$

$M_X(s) = \frac{\lambda}{s-\lambda}$ when, $s < \lambda$ ⓘ

- The formula for $M_X(s)$ is correct only if the integral $e^{(s-\lambda)x}$ decays as x increases which is the case if and only if $s < \lambda$, otherwise integral is infinite ;

Transform of Linear Function of a Random Variable :-

Random variable and $Y = aX + b$

Let, $M_X(s)$ be the transform associated with X.

$$M_Y(s) = E[\cancel{s}e^{sY}] = E[e^{s(ax+b)}] = E[e^{sax} \cdot e^{sb}]$$

$$= e^{sb} \cdot E[e^{sax}]$$

$$\therefore \boxed{M_Y(s) = e^{sb} \cdot M(as)}$$

- Transform associated with Standard r.v. :- $X \sim N(0,1)$

$$M_X(s) = E[e^{sX}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \quad \Big| \quad f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$\mu = 0$, $\sigma = 1$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} \cdot e^{-x^2/2} \, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x^2/2 - sx)} \cdot dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{x^2}{2}\right) + sx} \, dx$$

$$= \cancel{s}e^{+s^2/2} \int_{-\infty}^{\infty} e^{\left(-\frac{x^2}{2} + sx - \frac{s^2}{2}\right)} dx$$

$$= e^{s^2/2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-s)^2}{2}} dx}_{= 1 \text{ [total prob. = 1]}} \rightarrow \text{pdf of } X \sim N(s,1)$$

$$\Rightarrow \boxed{M_X(s) = e^{s^2/2}}$$

Transform Associated with a R.V. $X \sim N(\mu, \sigma^2)$:-

Soln: $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$\because$ We know the if $X \sim N(0,1)$ the, $\boxed{M_X(s) = e^{s^2/2}}$

and if $Y = aX + b$

then, $M_Y(s) = e^{sb} \cdot M_X(as)$

Let, $X = \sigma Y + \mu$

$$M_X(s) = e^{s\mu} \cdot M_X(\sigma s)$$

$$= e^{s\mu} \cdot e^{s^2\sigma^2/2}$$

$$M_X(s) = e^{(s^2\sigma^2/2 + \mu s)}$$

HW:- Compute the transform associated with -

(1) Geometric, (2) Poisson, (3) Binomial;

Note:- For C.R.V., we can get to moments by differentiation the transforms.

Ex - Show n^{th} moment of C.R.V. X, How to derive using transformation (n^{th} derivative at $s = 0$) $\rightarrow$ gives n^{th} moment;

For any random variable X, we have $M_X(s) = E[e^{sX}]$

So, $M_X(s)\big|_{at\ s=0} = E[e^0 X] = 1$

useful results

$$M_X(s) = \sum_{x=0}^{\infty} p(X=x)\, e^{sx}$$

Note: It is valid in this range only;

$$= p(X=0)\cdot e^{s\cdot 0} + \sum_{x=1}^{\infty} p(X=x)\, e^{sx}$$

As, $s \to -\infty$, all the terms e^{xs} with $x>0$, tend to 0, so we obtain,

$$\lim_{s\to -\infty} M_X(s) = P(X=0)$$ extracted pdf using MGF;

→ $Y = aX + b$
$M_Y(s) = e^{sb} M_Y(as)$

→ For $X \sim N(\mu, \sigma^2)$
$M_Y(s) = e^{\left(\frac{s^2\sigma^2}{2} + \mu s\right)}$

→ For $X \sim Geo(p)$
$M_Y(s) = \dfrac{pe^s}{1-qe^s}$

when $0<p\leq 1$

INVERSE TRANSFORM :-

Q) Let, $X_1 \sim N(\mu_1, \sigma_1^2)$; $X_2 \sim N(\mu_2, \sigma_2^2)$ and X_1, X_2 are indep.

Then, $Z = X_1 + X_2 \sim = ?$

Sol] $M_{X_1}(s) = e^{s^2(\sigma_1^2/2 + \mu_1 s)}$

$M_{X_2}(s) = e^{s^2(\sigma_2^2/2 + \mu_2 s)}$

$$M_{Z=X_1+X_2}(s) = E\left[e^{sX_1}\cdot e^{sX_2}\right] = E[e^{sX_1}]\cdot E[e^{sX_2}]$$

$$= e^{s^2(\sigma_1^2/2 + \mu_1 s)}\cdot e^{s^2(\sigma_2^2/2 + \mu_2 s)} = e^{s^2\left[\left(\frac{\sigma_1^2+\sigma_2^2}{2}\right) + s(\mu_1+\mu_2)\right]}$$

INVERSION OF TRANSFORMS :-

The transforms $M_Y(s)$ can be inverted i.e., $M_X(s)$ can be used to determine the prob. law that X follow;

(Inversion Property)

Theorem - The transforms $M_X(s)$ associates with a r.v. X uniquely determines the CDF of X, assuming $M_X(s)$ is finite $\forall\ s \in [-a, a]$ when 'a' is a positive number;

▷ Extract CDF of Geometric R.V. using transform of MGF :-

Q] We are told that the transform associated with a r.v. X is of the form $M_X(s) = \dfrac{pe^s}{1-qe^s}$ where, $q = 1-p$ and $0<p\leq 1$.

Find X follows what?

Sol] Given, $M_X(s) = \dfrac{pe^s}{1-qe^s}$

Here we use formula with $\alpha = (1-p)e^s$.

$1 + \alpha + \alpha^2 + \cdots = \dfrac{1}{1-\alpha}$ when $|\alpha|<1$

and for s sufficiently close to 0, so that $(1-p)e^s < 1$

$M_X(s) = pe^s[1 + qe^s + q^2e^{2s} + q^3e^{3s} + \dots]$

$= pe^s + pqe^{2s} + pq^2e^{3s} + \dots = \underbrace{p(1-p)^0}_{P(X=1)} e^s + \underbrace{p(1-p)}_{P(X=2)} e^{2s} + \underbrace{p(1-p)^2}_{P(X=3)} e^{3s}$

here we put parameter $s=0$ and get pmf of geometric;

➤ $X_1 \sim N(\mu_1, \sigma_1^2)$ and X_1, X_2 are indep. Figure out $X_1 + X_2 \sim ?$
$X_2 \sim N(\mu_2, \sigma_2^2)$

$\Rightarrow X_1 \sim N(\mu_1, \sigma_1^2) \Rightarrow M_{X_1}(s) = e^{(s^2\sigma_1^2/2 + \mu_1 s)}$
$X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow M_{X_2}(s) = e^{(s^2\sigma_2^2/2 + \mu_2 s)}$

Lemma:- Note: $M_{X_1+X_2}(t) = M_{X_1}(t) * M_{X_2}(t)$ when, X_1, X_2 are indep;

$M_{X_1+X_2}(s) = e^{s^2(\sigma_1^2+\sigma_2^2) + s(\mu_1+\mu_2)}$

proof out of syllabus.

X_1 O ↘ Same MGF : but this theorem gives uniquely (determine)
X_2 O ↗

Sum of of random variable :-

$X_1, X_2, X_3 \dots$ are iid r.v.s
then $X = X_1 + X_2 + \dots + X_n$

Sum of a Random number of Independent R.V.s :-

$X_1, X_2, \dots$ are independent and iid r.v.s. , and N is also r.v. such that, $N, X_1, X_2, \dots$ are also independent.

Let, $Y = X_1 + X_2 + \dots + X_N$

Common variance $= Var[X]$
Common expectation $= E[X]$
no. of r.v.s is also r.v. (N)

$E[Y] = E[N] \cdot E[X]$
↳ Proof by law of iterated expectation.

$Var[Y] = Var[Y|N=n] = Var[X_1 + X_2 + \dots + X_n | N=n]$
$= Var(X_1 + X_2 + \dots + X_n | N=n)$
$= Var(X_1 + X_2 + \dots + X_n) = Var(X_1) + \dots + Var(X_n)$
$= n\,Var(X)$ — iid (i.e., common var.)
↳ where, $Var(X) = Var(X_i)\ \forall i = 1, \dots, n$
$= n\,Var(X)$

$\therefore Var(Y|N) = N\,Var(X)$

Using Law of total variance,

$Var(Y) = E[Var(Y|N)] + Var(E[Y|N])$

$= E[\underbrace{N\,Var(X)}_{\text{constant } \forall N}] + Var(N \cdot \underbrace{E[X]}_{\text{constant}})$ — (derived) — from law of iterated expectation (see previous result)

$= Var(X) \cdot E[N] + (E[X])^2 Var(N)$

$\Rightarrow Var(Y) = E[N] \, Var(X) + (E[X])^2 \, Var(N)$

Transform Associated with Y :-

$M_Y(s) = E[e^{sY}] = E[E[e^{sY} | N]]$

Y is conditioned on X, so only we to bring X with law of iterated expectation.

The transform associates with Y conditioned on $N=n$

$E[e^{sY} | N=n] = E[e^{s(X_1+X_2+\cdots+X_N)} | N=n]$

$= E[e^{sX_1} \cdot e^{sX_2} \cdots e^{sX_n} | N=n]$

$= E[e^{sX_1} \cdot e^{sX_2} \cdots e^{sX_n}]$

Independent

$= E[e^{sX_1}] \cdot E[e^{sX_2}] \cdots E[e^{sX_n}]$

$= (M_X(s))^n$ —— (i) {as all have same MGF (iid r.v.s)}

$\Rightarrow \boxed{E[e^{sY} | N] = (M_Y(s))^N}$

$\therefore \; M_Y(s) = E[E[e^{sY} | N]] = E[(M_X(s))^N]$

$M_Y(s) = \sum_{n=0}^{\infty} (M_X(s))^N \cdot f_N(s)$ —— (ii)

When, $f_N(n)$ is the pmf of N ; —— (iii)

$M_N(s) = E[e^{sN}] = \sum_{n=0}^{\infty} e^{sn} f_N(n) \cdots$

Let us rework $(M_X(s))^n$ as, $(M_X(s))^n = e^{n \log(M_X(s))}$ —— (iv)

We can rewrite $M_Y(s) = \sum_{n=0}^{\infty} e^{n\log(M_X(s))} \cdot f_N(n)$ —— (v)

Comparing eqn (iii) & (v)

we get, $s = \log(M_X(s))$

We see that $M_Y(s)$ is obtained from $M_N(s)$ with 's' replace by $\log(M_X(s))$ or equivalently, e^s replaced by $M_X(s)$.

ex) $X_i \sim Geo(q)$ and $N \sim Geo(p)$ both iid ;

Figure out what $Y \sim$? $Y = X_1 + X_2$

Sol) $X_1 \sim Geo(q)$ $M_{X_1}(s) = \dfrac{q(e^s)}{1-(1-q)e^s}$

$N \sim Geo(p)$ $M_N(s) = \dfrac{p(e^s)}{1-(1-p)e^s}$

$M_{Y = X_1 + X_2}(s) = M_{X_1}(s) \cdot M_{X_2}(s)$

$= pqe^{2s}$

Sol:- Step1: Find MGF of N.

Step2: Г Transfor for each X_i

Step3: Replace with this $s = \log(M_X(s))$

Step4: Use theorem for proving the unique inverse ;

Class - Dt. 31/10/23

LIMIT THEOREM:-

$X_1, X_2, \dots$ a sequence of iid random variables; $E[X_i] = \mu$
$Var[X_i] = \sigma^2$

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

$E[S_n] = n\mu$, $Var(S_n) = n\sigma^2$

- $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$

$E[M_n] = \mu$ $\quad Var(M_n) = \frac{\sigma^2}{n}$

- $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, $E[Z_n] = 0$ $\quad Var(Z_n) = 1$ $\rightarrow$ Standard Normal!

Chebyshev:- $P(|X-\mu| \geq \varepsilon) \leq \sigma^2/\varepsilon^2$

Weak Law of Large Numbers:- (WLLN) Let, $X_1, X_2, \dots$ be a sequence of iid r.v.s. with mean μ and variance σ^2. Let, $M_n = \frac{X_1 + X_2 + \dots + \hat{X}_n}{n}$. Then, $E[M_n] = \mu$ and $var(M_n) = \frac{\sigma^2}{n}$. For every $\varepsilon > 0$, we have

$$P(|M_n - \mu| \geq \varepsilon) = P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2 n}$$

The RHS $\rightarrow 0$ as $n \rightarrow \infty$.

WLLN tells us that for larger n, the bulk of the distribution of M_n is concentrated around μ; (M_n converges to μ in probability)

Convergence in Probability:-

Convergence of a deterministic sequence - Let $x_1, x_2, \dots$ be a sequence of real numbers, and let a be another real number. We say that the sequence x_n converges to a, or $\lim_{n \to \infty} x_n = a$ if for every $\varepsilon > 0$, $\exists$ some n_0, s.t. $|x_n - a| \leq \varepsilon$ $\forall$ $n \geq n_0$.

Convergence in probability - Let $X_1, X_2, \dots$ be a sequence of r.v.s not necessarily independent, and let a be a real no. The sequence X_n converges to a in prob., if for every $\varepsilon > 0$, we have,

$$\lim_{n \to \infty} P(|X_n - a| \geq \varepsilon) = 0$$

suitably large n $\quad$ Confidence interval

$\Rightarrow P(|X_n - a| \geq \varepsilon) \leq \delta$ $\forall$ $n \geq n_0$ for every $\varepsilon > 0$ and $\delta > 0$.

Confidence level

- Why frequencies estimate probabilities?

Let A be an event corresponding to a probabilistic expt. and let, $p = P(A)$. We consider n ind. repetitions of the expt. and let,

$$X_i = \begin{cases} 1, & \text{if } i^{th} \text{ expt. throws up event } A \\ 0, & o/w. \end{cases}$$

$E[X_i] = P(X_i = 1) = p$

Consider the empirical frequency of A, $M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$

Let, $\mu = E[M_n] = \frac{1}{n}\sum_i E[X_i] = p$.

The WLLN says that, $P(|M_n - p| \geq \varepsilon) \to 0$ as $n \to \infty$

Convergence with Probability 1:- Let $X_1, X_2, \ldots$ be a sequence of r.v.s (not necessarily ind.) and let a be a real no. We say that the sequence X_n converges to a with prob. 1 (or almost surely (a.s.)) if, $P(\lim_{n\to\infty} X_n = a) = 1$

interpretation - Consider a sample space consisting of infinite sequences all of the prob. mass is concentrated on those sequences that converges to 'a'. This doesn't mean that other sequences are impossible, but they are extremely unlikely. (i.e., their total prob. tends to 0);

Ex: $X_1, X_2, \ldots$ be a sequence of iid r.v.s that are uniformly distributed in $[0,1]$. Let $Y_n = \min\{X_1, X_2, \ldots, X_n\}$. Show that Y_n converges to 0 a.s.

Ex: Consider a sequence of Y_n of non-ve r.v.s and suppose that $E[\sum_{n=1}^{\infty} Y_n] < \infty$. Show that Y_n converges to 0 a.s.;

Strong Law of Large Numbers (SLLN): Let, $X_1, X_2, \ldots$ be a sequence of iid r.v.s with mean μ. Then, the sequence of sample mean, $M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$ converges to μ, with prob. 1. ie.,

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right) = 1$$

Interpretation - Expt. generating the sample space Ω is infinitely long, and generates a sequence of values $x_1, x_2, \ldots$ conv. to r.v.s $X_1, X_2, \ldots$, So, think of Ω as a set of infinite sequences $(x_1, x_2, \ldots)$ of real nos; any such sequence is a possible outcome of the expt. Consider the set A of those sequences $(x_1, x_2, \ldots)$ whose long term average is μ. i.e.,

$$(x_1, x_2, \ldots) \in A \iff \lim_{n\to\infty} \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \mu$$

SLLN states that all of the prob. is concentrated on this particular subset of Ω.

WLLN says that $P(|M_n - \mu| \geq \varepsilon) \to 0$ as $n \to \infty$, is the probability of a "significant deviation" of M_n from μ goes -

to zero as $n \to \infty$. But, for a finite n, the prob. can be +ve, and it is possible that even if not very frequently, but M_n deviates significantly from μ. WLLN doesn't provide any guarantee on the # such deviations but SLLN does.

SLLN implies that for any given $\varepsilon > 0$,

$$P(\#(|M_n - \mu| \geq \varepsilon) = \infty) = 0 .$$

Extra: $\boxed{\text{Thm 1:-}}$

Let $Y_1, Y_2, \ldots, Y_n$ be <u>non negative</u> sequence of random variables. Given that, $E\left[\sum_{i=1}^{\infty} Y_i\right] < \infty$, Y_n converges to 0 a.s.

$\rightarrow E\left[\sum_{i=1}^{\infty} Y_i\right] = \sum_{i=1}^{\infty} E[Y_i]$. (Monotone Convergence Theorem)

$\therefore \sum_{i=1}^{\infty} E[Y_i] < \infty \Rightarrow \sum_{i=1}^{\infty} Y_i$ is always finite.

Series Converges $\Rightarrow$ Sequence Converges to 0 ;

if not let K

$\downarrow$

S_n, S_{n+1}, S_{n+2}

$\downarrow \quad \downarrow \quad \downarrow$

$l \quad l+K \quad l+2K \leftrightarrow$ if we choose $\varepsilon < K$ then series never converges!

<u>Monotone Conv. Theorem:-</u>

Any non decreasing sequence that is bounded above converges;

<u>ex1</u> Consider a sequence of Bernoulli r.v.s, $X_1, X_2, \ldots, X_n$ and let $p_i = \Pr(X_i = 1)$. Assume that, $\sum_{n=1}^{\infty} p_n < \infty$. Then, <u>no. of successes</u> is finite with prob. 1.

$$\sum_{i=1}^{n} X_i \qquad E\left[\sum_{i=1}^{\infty} X_i\right] = \sum_{i=1}^{\infty} E[X_i] = \sum_{i=1}^{\infty} p_i < \infty$$

$\rightarrow \sum_{i=1}^{\infty} X_i$ is finite with probability 1 ;

<u>Thm 2:</u> Let, $X_1, X_2, \ldots$ be a sequence of iid r.v.s & assume $E[X^4] < \infty$. Prove strong LLN ;

$$\hookrightarrow P\left(\lim_{n \to \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right) = 1$$

$$\hookrightarrow P\left(\lim_{n \to \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu = 0\right)$$

Assume, $E[X_i] = 0$

$$X_i' = X_i - \mu$$

$$E\left[\frac{(X_1+X_2+\cdots+X_n)^4}{n^4}\right] = \frac{1}{n^4} E\left[(X_1+X_2+\cdots+X_n)^4\right] \quad (1)$$

$(X_1+X_2+X_3+\cdots+X_n)^4$

$E[X_{i_1}] = E[X_{i_1}^3] = \cdots = 0$

$X_{i_1} X_{i_2} X_{i_3} X_{i_4} \to 0$

$X_{i_1} X_{i_1} X_{i_2} X_{i_3} \to 0$

$X_{i_1} X_{i_1} X_{i_2} X_{i_2}$ ✓

$X_{i_1} X_{i_1} X_{i_1} X_{i_2} \to 0$

$X_{i_1} X_{i_1} X_{i_1} X_{i_1}$ ✓

$X_{i_1} X_{i_2} X_{i_1} X_{i_2}$

$X_{i_2} X_{i_1} X_{i_2} X_{i_1}$

$X_{i_1} X_{i_1} X_{i_2} X_{i_2}$

$X_{i_1} X_{i_2} X_{i_2} X_{i_1}$

$\frac{{}^4C_2}{2}$

from (1), $= \frac{1}{n^4} \cdot n E[X_i^4] + \underbrace{\frac{3n(n-1)}{2n^4} E[X_{i_1}^2 X_{i_2}^2]}$

$\frac{3\,(n-1)}{2\,n^3} E[X_{i_1}^4]$

we know,

$$\boxed{\begin{aligned} &X^2+Y^2 \geq 2XY \\ &\Rightarrow \frac{X^2+Y^2}{2} \geq XY \\ &\Rightarrow E[X_i^2 X_j^2] \leq E\left[\frac{X_i^4+X_j^4}{2}\right] \\ &\qquad = E[X_i^4] \end{aligned}}$$

$= \frac{n(3n-2)}{n^4} E[X_i^4]$

$\leq \underset{\substack{\downarrow \\ 0}}{\frac{3}{n^2}} \underbrace{E[X_i^4]}_{<\infty}$ as $n \to \infty$

$$\Rightarrow \sum_{i=1}^{\infty} E\left[\frac{(X_1+X_2+\cdots+X_n)^4}{n^4}\right]$$

$$= \sum_{i=1}^{\infty} \frac{3}{n^2} \underbrace{E[X_i^4]} \Rightarrow K \text{ (say)}$$

$\left[\frac{1}{n} \text{ series not conv.} \atop \text{but } \frac{1}{n^2} \text{ conv.}\right]$

$$= 3K \cdot \underbrace{\sum_{i=1}^{\infty} \frac{1}{n^2}}_{<\infty}$$ finite then it conv. zero!

$$\therefore \Pr\left(\lim_{n\to\infty} \frac{(X_1+X_2+\cdots+X_n)^4}{n^4} = 0\right) = 1 ; \quad \square$$

$$\Downarrow$$

$$\Pr\left(\lim_{n\to\infty} \frac{X_1+X_2+\cdots+X_n}{n} = 0\right) = 1 ;$$

Class- Dt. 04/11/23

A bernoulli process is a sequence $X_1, X_2, \ldots$ of ind. bernoulli r.v.s X_i's with, $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$; || Fresh Start

T: time (# trials) till the first success.

$T \sim Geo(p)$

$P(T = n+K \mid T > K) = P(T = n)$

$\Rightarrow P(\underbrace{T-K}_{} = n \mid T > K) = P(T = n)$

after K failures, # trials to get first success (fresh start property)

Independence Properties :-

- For any given time n, the sequence of r.v.s $X_{n+1}, X_{n+2}, \ldots$ is also a Bernoulli process and is independent of $X_1, X_2, \ldots, X_n$.
- Let K be a given time and T : time of first success after time K., Then $T-K$ has a geometric distribution with parameter p, and is independent of the r.v.s $X_1, X_2, \ldots, X_n$.

Ex: Let N be the first time that a success has happened immediately after another success ; i.e., N is the first i for which $X_i = X_{i-1} = 1$. What is the probability that there are no successes in the next two trials. i.e.,

$P(X_{N+2} = X_{N+1} = 0) = ?$ .... $\overset{N}{S}\underset{i}{S}FF$

$\Rightarrow (1-p)^2$ (By fresh start property) ;

Interarrival Times :-

Y_K : time / # trials till the K^{th} success.

T_K : K^{th} interarrival time.

Is defined as $T_1 = Y_1$; $T_K = Y_K - Y_{K-1}$, $K = 2, 3, 4, \ldots$

$$T_2 = Y_2 - Y_1 \quad \Rightarrow Y_2 = T_1 + T_2$$

$$Y_3 = T_1 + T_2 + T_3$$

$$\vdots$$

$$Y_K = T_1 + T_2 + \cdots + T_K$$

$T_1 \sim Geo(p)$; $T_i \sim Geo(p)$ (by fresh start property)

- $E[Y_K] = \frac{K}{p}$ $\quad Var(Y_K) = K\frac{(1-p)}{p^2}$

Alternate way of defining the Bernoulli's Process :-

- Start with a sequence of ind. geo. r.v.s $T_1, T_2, T_3, \dots$ with common parameter p, and let them be the interarrival times.
- Record a success/arrival at times,

$T_1, T_1+T_2, T_1+T_2+T_3, \dots$

PMF of Y_K: $P(Y_K=t) = \binom{t-1}{K-1} \cdot p^{K-1} \cdot (1-p)^{t-1-(K-1)} \cdot p$

$$\underbrace{\cdots\cdots t-1 \cdots\cdots}_{K-1} \underset{p}{\overset{t}{\textcircled{S}}} \qquad = \binom{t-1}{K-1} p^K (1-p)^{t-K}$$

ex: Jobs are coming, we have to formulate a sequence of X_i's for accepting jobs and Y_i's for rejecting jobs.

Accept :- $X_{n_1}, X_{n_2}, \dots \rightarrow \text{Ber}(pq)$ Parent Bernoulli

Reject :- $X_{m_1}, X_{m_2}, \dots \rightarrow \text{Ber}(p(1-q))$

→ Merging / splitting Bernoulli process → A Bernoulli process;

MARTINGALES:- (conditional expectation recap...)

- $E[X|Y=y] = \sum_x x P(X=x | Y=y)$
- $E[E[X|Y]] = E[X]$
- For any two r.v.s X & Y, $E[Y \cdot E[X|Y]] = E[XY]$

proof: $E[Y \cdot E[X|Y]] = \sum_y P(Y=y) \cdot y \cdot \underbrace{E[X|Y=y]}$

$= \sum_y P(Y=y) \cdot y \cdot \sum_x x P(X=x|Y=y)$

$= \sum_x \sum_y xy \cdot \{P(X=x|Y=y) \cdot P(Y=y)\}$

$= \sum_x \sum_y xy \, P(X=x, Y=y)$

$= E[XY]$

- $E[X|Z] = E[E[X|Y,Z] | Z]$

Martingale: $\overset{\text{defn}}{\text{A}}$ sequence of R.V.s $Z_0, Z_1, \dots$ is a martingale w.r.t. a sequence $X_0, X_1, \dots$ if $\forall n \geq 0$, the following condn

- hold :- (For a sequence dependent on another seq. of R.V.s)

- Z_n is a function of $X_0, X_1, \dots, X_n$;
- $E[|Z_n|] < \infty$;
- $E[Z_{n+1} | X_0, \dots, X_n] = Z_n$;

restriction on Z_{n+1} so that it not blows up.

Z_n

$\underbrace{X_0, X_1, \dots, X_n}_{}, \underbrace{X_{n+1}}_{}$

doesn't vary Z_{n+1} too much than Z_n;

Z_{n+1} varies from Z_n but not in terms of expectation

defn: A sequence of r.v.s $Z_0, Z_1, \dots$ is called a Martingale when it is a Martingale w.r.t itself. i.e.,

$E[|Z_n|] < \infty$ and $E[Z_{n+1} | Z_0, \dots, Z_n] = Z_n$

Lemma:- Let $Z_0, Z_1, \dots$ be a Martingale sequence. Then $\forall i$,

$E[Z_i] = E[Z_0]$;

$\mapsto E[Z_1 | Z_0] = Z_0 \Rightarrow E[E[Z_1|Z_0]] = E[Z_0] \Rightarrow E[Z_1] = E[Z_0]$

ex:① A gambler plays a sequence of fair games. Let,

X_i: amount the gambler wins at the i^{th} game.

Z_i: gambler's total win at the end of the ith game.

Each game is fair $E[X_i] = 0$. Whether, Z_i is a Martingale?

$\rightarrow E[Z_{i+1} | X_1, X_2, \dots, X_i] = Z_i + \cancel{E[X_{i+1}]}^{\,0}$

$= Z_i$

ex② Let $X_1, X_2, \dots$ be ind r.v.s with zero mean and let $Z_n = \sum_{i=1}^{n} X_i$. Then, $\{Z_n | n \geq 1\}$ is a Martingale;

ex③: Let $X_1, X_2, \dots$ be ind rvs with mean 1 and let $Z_n = \prod_{i=1}^{n} X_i$

Then, $\{Z_n | n \geq 1\}$ is a Martingale?

ex④: Consider Branching Process and let X_n denote the size of the nth generation. If m is the mean number of offspring per individual, then $\{Z_n | n \geq 1\}$ is a Martingale when, $Z_n = X_n / m^n$;

DOOB MARTINGALE :-

Let $X_0, X_1, \ldots$ be a sequence of r.v.s and let Y be a r.v. with $E[|Y|] < \infty$, Y will depend on $X_0, X_1, \ldots, X_n$. Then, $Z_i = E[Y \mid X_0, \ldots, X_i]$, $i = 0, 1, \ldots, n$ gives a Martingale w.r.t $X_0, X_1, \ldots$ Since, $E[Z_{i+1} \mid X_0, X_1, \ldots, X_i] = E[\underbrace{E[Y \mid X_0, X_1, \ldots, X_{i+1}]} \mid X_0, X_1, \ldots, X_i]$

"Filtration Sequence"

$= E[Y \mid X_0, X_1, \ldots, X_i]$

$= Z_i$ (law of conditional iterated expectation).

Doob Martingales for "Random Graphs" :-

Random Graph: $G(n,p)$: a random graph on n vertices. There will be an edge between two vertices with a prob. of p;

Let, $X_j = \begin{cases} 1, & \text{if there is an edge in the } j^{th} \text{ slot} ; 1 \leq j \leq \binom{n}{2} = m \\ 0, & \text{o.w.} \end{cases}$

Consider any finite valued function defined over graphs. Let $F(G)$ be the Maximum Independent Set in G. Let, $Z_0 = E[F(G)]$ and $Z_i = E[F(G) \mid X_1, \ldots, X_i]$, $i = 1, \ldots, m$

The sequence Z_i is a Doob Martingale. This is known as the edge exposure Martingale;

~ly (Similarly), we could expose the set of edges connected to a given vertex, one vertex at a time. The vertices ~~and~~ are numbered $1, \ldots, n$. and let G_i : subgraph of G induced by the first i vertices.

Set, $Z_0 = E[F(G)]$ and $Z_i = E[F(G) \mid G_1, G_2, \ldots, G_i]$ $i = 1, 2, \ldots, n$. Z_i's form a Doob Martingale known as the vertex exposure Martingale;

Date
14 Nov 23

* Martingale Process *

↳ fair batting game.

* Stopping time :-

Q batting Game

before Game It is decided :- I will stop after n many steps it is fixed

Q:- If I will loose 100₹ or win 100₹

Lemma :- If the sequence $Z_0, Z_1, \ldots Z_n$ is a martingale w.r.t $X_0, X_1 \ldots X_n$ then $E[Z_n] = E[Z_0]$

Proof :- $E[Z_{i+1} \mid X_0, X_1 \ldots X_i] = Z_i$ (by definition of Martingale)

taking Expectation on both side,

$$\Rightarrow E\left[E[Z_{i+1} \mid X_0, X_1, \ldots X_i] \right] = E[Z_i]$$

$$\Rightarrow E[Z_{i+1}] = E[Z_i]$$

Repeat this, to get $\boxed{E[Z_0] = E[Z_n]}$

meaning If no. of Game played is predetermined then Expectation of winning is Zero.

Stopping time :-

Definition : A non-negative integer valued random variable T is a stopping time for the sequence $\{Z_i : i \geq 0\}$ if the prob. of the event $\{T = n\}$ is independent of the variables

$$\{Z_{n+j} \mid Z_1, Z_2, \ldots Z_n, j \geq 1\}$$ (ie the variable $Z_{n+1}, Z_{n+2} \ldots$ conditioned on the values $Z_1, Z_2 \ldots Z_n$)

Intution :- if I stop at $t = 10$, then it does not ~~matter with~~ matter what will happen after $t = 11, 12, \ldots$

Martingale Stopping Theorem :- If $Z_0, Z_1 \ldots$ is a martingale w.r.t $X_1, X_2 \ldots$ and if T is a stopping time for $X_1, X_2 \ldots$ then $E[Z_T] = E[Z_0]$ whenever one of the follow holds :-

T is random variable ; So this is dependent on T. (Z)

(i) Z_i are bounded i.e. $\forall i$ $|Z_i| \leq C$; where C is a constant

(ii) T is bounded

(iii) $E[T] < \infty$, then there is a constant C, such that

$$E\left[\,|Z_{i+1} - Z_i| \;\middle|\; X_1, X_2, \ldots X_i\right] < C$$

Gambler's Ruin :-

Z_i = Sum of Game till now = $\sum X_i$

$X_i \to$ ~~Game at ith~~ amount won at ith Game

$Z_i \to$ total amount won after the ith Game.

and $Z_0 = 0$

Stopping time:-

The player quits the game when the player either loose l_1 amount of money $= ₹50$ or, l_2 amount of money win ₹100

Q. What is the probability that the player wins l_2 amount of money $= ₹100$ before losing $l_1 = ₹50$?

Solution:- Let T : first time the player has either won l_2 or lost l_1.

So, T is a stopping time for $X_1, X_2, \ldots$

The sequence $Z_0, Z_1 \ldots$ is a martingale and Z_i's are bounded, so martingale stopping theorem works, and we have

$$E[Z_T] = E[Z_0] = 0$$

Let prob. that the player wins l_2 amount before loosing $l_1 = q$

$\therefore$ $E[Z_T] = l_2 q - l_1(1-q)$. (from def.)

and $E[Z_T] = 0$

$$0 = l_2 q - l_1 + l_1 q$$

$$\boxed{q = \frac{l_1}{l_1 + l_2}} = \frac{50}{150} = \frac{1}{3} = 33\%$$

Q. **WALD'S EQUATION**:— Let $X_1, X_2, \ldots$ be a non-negative iid random variables with distribution X. Let T : stopping time for sequence $X_1, X_2, \ldots$

If T and X have bounded expectation then

$$E\left[\sum_{i=1}^{T} X_i\right] = E[T] \cdot E[X]$$

[we did this earlier ~~proofs~~]

Proof:- For $i \geq 1$, Let

$$Z_i = \sum_{j=1}^{i} (X_j - E[X])$$

Exercise:- The sequence Z_i forms a martingale w.r.t $X_1, X_2, \ldots$. (Proof)

$$E[Z_1] = \sum_{j=1}^{i} (X_1 - E[X]) = 0$$

and $E[T] < \infty$, $E[X] < \infty$ } bounded ~~[illegible]~~ Expectation

Now

$$Z_{i+1} = \sum_{j=1}^{i+1} (X_j - E[X])$$

$$Z_i = \sum_{j=1}^{i} X_j - E[X]$$

$$\therefore \quad Z_{i+1} - Z_i = (X_{i+1} - E[X])$$

$$E\left[|Z_{i+1} - Z_i| \mid X_1, X_2 \ldots X_i\right] = E\left[X_{i+1} - E(X)\right] \leq 2 \cdot E[X] < \infty$$

ie. is bounded

Now

$$Z_T = \sum_{i=1}^{T} (X_j - E[X])$$

$$E[Z_T] = \sum_{j=1}^{T} (X_j - E[X]) = 0 \qquad \left[\text{b/c } E[Z_T] = E[Z_0] = 0\right]$$

$$\Rightarrow E[Z_1] = \sum_{j=1}^{T} (X_j - E[X]) = 0$$

$$= \quad E\left[\sum_{j=1}^{T} X_j\right] - T \cdot \sum_{j=1}^{T} - E[X] = 0$$

$$\sum_{i=1}^{T} X_i = T \cdot E[X]$$

taking expectation on both sides

$$E\left[\sum_{i=1}^{T} X_i\right] = E[T] \cdot E[X]$$ proved

Simpler definition of stopping time,

* **Stopping time :-** Let $Z_0, Z_1, \ldots$ be a sequence of iid r.v.s, A non-negative integer valued rv T is a stopping time for the sequence if the event $\{T = n\}$ is independent of $Z_{n+1}, Z_{n+2}, \ldots$

Exercise :-

fair dice roll

if no. comes up → 3 ⇒ roll 3 times

→ 4 ⇒ roll 4 th

i.e. rolling a dice for random variable

Q. what will be sum of dice

$$E\left[\sum_{i=1}^{T} X_i\right] = E[T] \cdot E[X]$$

$$= 7/2 \cdot 7/2 = \left(\frac{7}{2}\right)^2$$

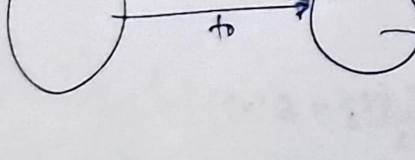
Q. How to decide a Data str for a prob ?

Ans :- (i) data is dynamic or static ?

(ii) what operation you want to perform

Q. When we use Hashing ?

A ⇒ maps the big universe small universe



⇒ use it for searching

Q. choose a hash function uniformly at random for Univer Hash family what is meaning ?

A ⇒ Choosing coefficient of polynom from a range so that $h(x)$ becomes random.

A> $h : \mathcal{U} \rightarrow [n]$

universe

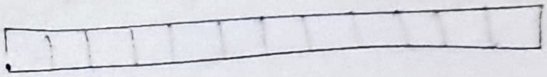
$h(x) : (ax+b) \bmod n$

choose randomly a, b from a range $[n]$

> storing coeffi'n of polynomial take $O(\log n)$ ~~time~~ space

Hashing is great example when search $O(1)$ time

QuickSort :-

arr[n] : 

Input : array of size n. (the number in the array are distinct)

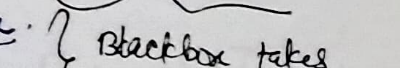
step1:- Pick a pivot = 'a'

it can be 1^{st} el
- last el
- mid el
- random elmt

step2:- ~~Spli~~ Partition the array into two parts

not sorted

$x<a$ $x>a$



a

$X_{<a} = \{x \in A \mid x<a\}$

$X_{>a} = \{x \in A \mid x>a\}$

a is correct position

beauty of Quicksort :- partition in inplace.

Blackbox takes $O(n)$ comparison to split

Step3:- Recursively partition the array $X_{<a}$ and $X_{>a}$

Qsort ($X_{<a}$)

Qsort ($X_{>a}$)

if Good part $T(n) = 2T(n/2) + O(n)$

Prob. Analysis of Random Algo.

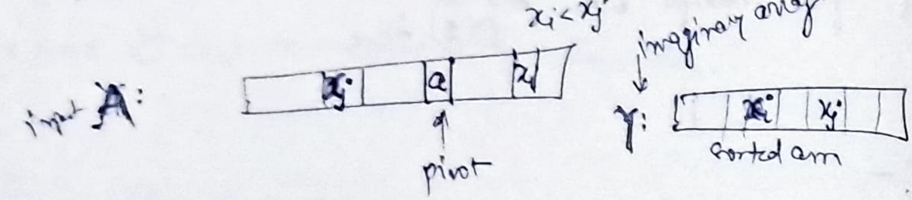
Randomized Quicksort :- Pick a pivot uniformly at random

Vs

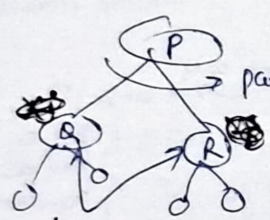
Prob. analysis of deterministic algo. (input could be)

Q. How to Analyse

→ Given input A:
I think an imaginary array Y, when elements in Y are written in increasing order



two no. in array can't be compare more than once



We never compare element from this set (Q) to that set (R).

$$X_{ij} = \begin{cases} 1 : \text{if } x_i \text{ and } x_j \text{ were compared} \\ 0 ; \text{ o/w} \end{cases}$$

$\therefore$ total no. of comparison $= \displaystyle\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$

$$\therefore E[\#\text{ Comparison}] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(X_{ij}=1)$$

Problem 10 :- Consider 2 sequence of letters (a-z), A and B, stored in arrays.

(a) Write a program to find the no. of (possibly overlapping) occurrences of the sequence B in A.

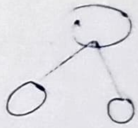
Q. How to find $P(X_{ij} = 1)$

A. ~~Imagine an : r: sorted array~~

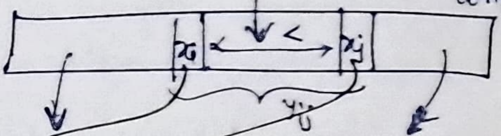
~~this span $(j-i)+1$~~

~~any elmt in this range will be selected as pivot then x_i and x_j never be compared~~

$$P(X_{ij} = 1) = \frac{2}{j-i+1}$$



if pivot selected in this range then they never compared



if x_i and x_j will be in same partition, then only be compared.

if pivot is selected in this range then it won't affect x_i and x_j comparison.

out of two possibility of pivot $= (x_i \text{ or } x_j)$ will be selected from subarray index $(j-i)+1$

$$E[\#\text{comparation}] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{(j-i+1)}$$