6/1/20

PROBABILITY THEORY AND RANDOM PROCESSES

Text book:- Probability, Random variables and Stochastic processes
(A. Papoulis, S.U. Pillai)

Office Room: 315 (Science block)

Exam pattern:-

Quiz-I:- 10 marks
M.S :- 30 marks
Quiz-II:- 10 marks
E.S :- 50 marks (10-12 marks from M.S)

* The theory of probability deals with phenomena which are random in nature which under some statistical experiments yields outcome indicating some pattern about the quantity of interest. Statistical experiment is also called random experiment.

* Random experiment:-
  - All possible outcomes are known in advance
  - Any particular trial will yield an outcome that is not known in advance
  - Experiment can be repeated under identical conditions.

* All possible outcomes are stored in the sample space. (S or Ω)
It is a collection of all outcomes of a random expt.

Ex:- Toss a coin once $S = \{H, T\}$

Toss a coin ten times $S = \{ (e_1, e_2, \dots e_{10}) : e_i = H \text{ or } T, \; i = 1 \text{ to } 10 \}$

The same experiment may have many sample spaces. It is upto us as to which one we choose to work with.

Eg:- life time of a machine :-

$S_1 = \{ t : t \geq 0 \}$ $t \to$ time in years

$S_2 = \{ t : 0 \leq t \leq N \}$

$S_3 = \{ N : n = 1, 2, 3 \dots$ in weeks $\}$

* An Event is some subset of a sample space.

Eg: In the previous case, S_2 is an event of S_1.

We can define further events from one or more existing events :-

Let A and B are events in S, then

- $A \cup B \to$ occurance of atleast one of events A and B.

Generalizing, Let $A_1, A_2, \dots A_n$ be events in S, then,

$\bigcup_{i=1}^{n} A_i \to$ occurance of atleast one of $A_1, A_2 \dots A_n$.

- $A \cap B \to$ Simultaneous occurance of events A and B

- $A^c \to$ Occurance of the complementary event.

and other elementary operations.

* Once we have events of our interest we can assign probabilities to them.

* Classical definition of probability (Laplace 1812) :-

Consider a random event whose outcomes are stored in the sample space S which contain n points. Then probability of an event A with m favourable outcomes is given by :-

$$P(A) = \frac{m}{n}$$

Assumptions here :-

- We have only finite no. of total outcomes (n).
- All n outcomes are equally likely to occur.

(This definition is circular in nature as we are considering probability of events (equally likely) to define the term probability).

* History :-
- The game of chances lead to the origin of theory of probability.
- It can be traced back to the 17th century.
- Owners of gambling houses requested mathematicians like Fermat, Pascal, DeMere to develop this concept.

* Probability has found wide application in many fields as tools of probability and statistics can be used to explain and understand variations involved.

* Relative frequency definition of probability (Empirical definition):

Consider a random event with some sample space S. Let the experiment be conducted n number of times, and n_A be the number of outcomes for event A (in favour of), then

$$\text{relative frequency of } A = \frac{n_A}{n}$$

We keep conducting the experiment many number of times, the stabilization of this frequency for long sequence of trials for a random event is known as statistical

So, $p(A) = \frac{n_A}{n}$ as $n \to \infty$.

Here we are not assuming finite no. of trials or equally likely outcome for events, but it is not a practical concept as $n \to \infty$.

We may have functions like,

$\frac{\sqrt{n}}{n}$, we know as $n \to \infty$ $\frac{\sqrt{n}}{n} \to 0$

$\Rightarrow p(A) = 0$.

$\frac{n - \sqrt{n}}{n}$ as $n \to \infty$, tends to $1 \Rightarrow p(A) = 1$

The value the func. converges to as $n \to \infty$, is the probability.

* Axiomatic definition of probability:

This is the most accepted definition of probability.

Let S be a sample space and $P(A)$ denote the probability of event A in S, then,

Axiom I :- $P(A) \geq 0 \ \forall A \in S$

Axiom II :- $P(S) = 1$

Axiom III :- Let A and B be two disjoint events in S, then,

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

Axiom III* :- Let $A_1, A_2, \dots A_n$ be pairwise disjoint events in S, then,

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

Pairwise disjoint ⇒

- $P(\phi) = 0$ Proof:-

$$\phi \cup \phi^c = S \Rightarrow P(\phi \cup \phi^c) = P(S)$$

$$\Rightarrow P(\phi) + P(\phi^c) = P(S)$$

$$\Rightarrow P(\phi) + P(S) = P(S)$$

- $P(A^c) = 1 - P(A)$ for $A \in S$

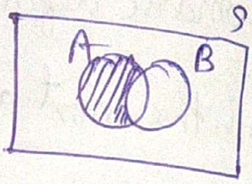
  (also stated as $A \cup A^c = S$)

- $0 \leq P(A) \leq 1$

- If $A \subseteq B$ then $P(A) \leq P(B)$

- $P(B^c \cap A) = P(\overset{A}{B}) - P(A \cap B)$

  (proved from $A = (A \cap B) \cup (A \cap B^c)$)



- Addition formula for probability :-

  Let A and B be two events in S. Then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

  proof:-
  $$A \cup B = \underbrace{(A \cap B^c) \cup (A \cap B) \cup (A^c \cap B)}_{\text{three disjoint events} \Rightarrow}$$

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

- General addition formula of probability :-

  For n events, $A_1, A_2, \ldots A_n$,

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

(Can be proved using PMI). It is a very quite complicated formula and a much simplified version can be used as,

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

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

\# Three popular options for a car model are :-

A :- automatic B :- new age engine C : Air conditioning

we have the following data,

$P(A) = 0.7$ $P(B) = 0.75$ $P(C) = 0.8$ $P(A \cup B) = 0.8$

$P(A \cup C) = 0.85$ $P(B \cup C) = 0.9$ $P(A \cup B \cup C) = 0.95$

Find the prob. that a buyer chooses exactly one of the option.

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

\# Toss a fair die twice. We define the events,

A : { atleast one six }

$P(A) = \underline{\underline{11/36}}$ $\left(\frac{6}{\ } \frac{\ }{5} \quad \frac{5}{\ } \frac{6}{\ } \quad \frac{6}{\ } \frac{6}{\ } \right)$

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

B : { Sum of faces ≥ 10 }

When this event is known, what is the $P(A) = ?$

$\Rightarrow$ 6 cases in total where sum ≥ 10 (∴ sample space has to be restricted)

Now, atleast one six $\Rightarrow$ (6 5, 6 4, 6 6, 5 6, 4 6)

$= \underline{\underline{5/6}}$

* Conditional probability:-

Consider a sample space S where let A and B are two events then conditional probability of A given that event B has already occured is defined as

$$P(A|B) = \frac{P(A\cap B)}{P(B)}, \quad P(B) > 0$$

Now, if $A \subseteq B$, $P(A|B) = \frac{P(A)}{P(B)}$ which is $> P(A)$ since $P(B) < 1$

and if $B \subseteq A$, $P(A|B) = 1$.

Axiom I:- $P(A|B) \geq 0 \quad A \in S$

Axiom II:- $P(S|B) = 1$

Axiom III:- If A and C are disjoint events in S then

$$P(A \cup C|B) = P(A|B) + P(C|B)$$

* Multiplication Rule:-

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

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

Generalizing,

let $A_1, A_2, \dots A_n$ be n events in S, then,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2) \dots P\left(A_n \middle| \bigcap_{i=1}^{n-1} A_i\right)$$

* Independent events :- Two events A and B are independent when $P(A\cap B) = P(A) \cdot P(B)$

Disjoint events with positive probabilities ($P(A)>0$ and $P(B)>0$) can never be independent.

Toss a coin 3 times. $S = \{HHH, HHT, HTH, HTT, TTH, THT, THH, TTT\}$

A : head on first coin
B : " " 2^{nd} coin
C : head on 3^{rd} coin

Are A, B, C independent events?

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

In order to verify whether A, B, C are independent events we need to check,

$$\left.\begin{array}{l} P(A\cap B) = P(A)P(B) \\ P(B\cap C) = P(B)P(C) \\ P(C\cap A) = P(C)P(A) \end{array}\right\} \text{pairwise exclusive}$$

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

then they are mutually exclusive in nature.

Here $P(A\cap B) = \frac{1}{4} = P(A)P(B)$

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

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

$P(A\cap B\cap C) = P(A)P(B)P(C) = 1/8$

$\Rightarrow$ these events are mutually exclusive.

* Now consider n events,

$A_1, A_2, \dots A_n$, then to check their mutual exclusivity we have the following condition,

$$P(A_i \cap A_j) = P(A_i)\,P(A_j) \quad i<j=1,2,\dots n$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)\,P(A_j)\,P(A_k) \quad i<j<k=1,2,3\dots n$$

... and so on

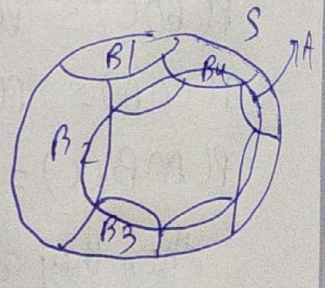
=) total no. of conditions to check are

$$^nC_2 + {^nC_3} + {^nC_4} + \dots {^nC_n} = 2^n - n - 1$$

* Total probability result :-

Let $B_1, B_2, \dots B_n$ be n events of a sample space which form a partition of the given sample space which also contains the event A. Then,

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



Proof:- Since $B_1, B_2, B_3 \dots B_n$ form a clear partition of S, hence these events are pairwise disjoint and they satisfy

$$\bigcup_{i=1}^{n} B_i = S \qquad \Rightarrow B_i \cap B_j = \phi$$
$$(i<j=1,2,3\dots n)$$

we have

$$A = (A\cap B_1) \cup (A \cap B_2) \cup (A\cap B_3) \dots \cup (A \cap B_n)$$

Since A is a part combination of one or more of B_n.

Also note that,

$A \cap B_i$ and $A \cap B_j$ are pairwise disjoint $(i, j = 1, 2, \dots n)$

$$\Rightarrow P(A) = P((A \cap B_1) \cup (A \cap B_2) \dots \cup (A \cap B_n))$$

$$= P(A \cap B_1) + P(A \cap B_2) \dots + P(A \cap B_n) \text{ (as it is pairwise disjoint)}$$

$$= P(B_1)\, P(A/B_1) + P(B_2)\, P(A/B_2) + \dots P(B_n)\, P(A/B_n)$$

* Bayes Theorem:-

Let $B_1, B_2, \dots B_n$ be a partition of S and A be any event in S. Then,

$$P(B_i/A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(B_i)\, P(A/B_i)}{\sum_{j=1}^{n} P(B_j)\, P(A/B_j)} \quad j = 1, 2, 3 \dots n$$

$P(B_i) \rightarrow$ a priori event/probability

$P(B_i/A) \rightarrow$ posterior probability.

Here, we have $P(B_i)$ and using the data of a new event A, we again update our knowledge of the event B_i by finding $P(B_i/A)$. It is also called inverse probability.

(Also, $\sum_{\text{all } i} P(B_i/A) = 1$)

Let three vendors A, B, C supply same product to IIT Patna. Suppose vendors supply 35%, 35%, 30% of the product. Also it is known that 8%, 10% and 5% of these products tend to be defective. Suppose an item is tested at random,

what is the probability that it is defective?

Given that it is defective, what is the probability that it was supplied by A, B, C respectively.

$\Rightarrow$ $P(\text{defective}) = P(\text{vendor } 1) \cdot P(\text{defective}/\text{vendor } 1) + P(V2) \cdot P(\text{def.}/V2) + P(V3) \cdot P(\text{def}/V3)$

$$= \frac{35}{100} \cdot \frac{8/100}{35/100} + \cancel{P(V_2)} \cdot \frac{P(\text{def} \cap V_2)}{\cancel{P(V_2)}} + \cancel{P(V_3)} \cdot \frac{P(\text{def} \cap V_3)}{\cancel{P(V_3)}}$$

$$= \frac{8}{100} + \frac{10}{100} + \frac{5}{100} = \underline{0.23}.$$

A binary communication channel carries data as one of two types of signals 0 and 1. Due to noise, a transmitted zero is sometime recieved as 1 and vice versa. For a given channel, assume that a prob. of 0.94 that a transmitted 0 is correctly recieved and a prob. of 0.91 that a transmitted 1 is correctly recieved. Father assume that prob. of transmitting a 0 is 0.45. If signal is sent, find the probability that

(i) a 1 is recieved

(ii) a 1 was transmitted given that a 1 was recieved

(iii) an error occurs in transmission.

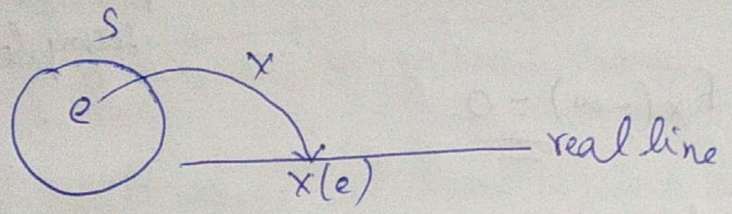
(i) 0 $\to$ transmit, 1 $\to$ recieved + 1 $\to$ transmit, 1 $\to$ recieved

$\Rightarrow$ $0.45 \times 0.06 + 0.55 \times 0.91$

* Random variables:-

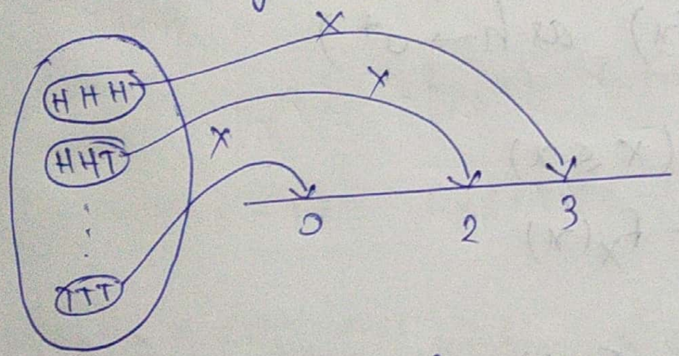
consider a random experiment with some sample space S.

A random variable X is a function defined over S which associates each outcome e in S to a real number.



\# Toss a coin 3 times and

X :- No. of heads in the outcome



R_X : All possible values of X

$\Rightarrow$ In the above case, $R_X : \{0, 1, 2, 3\}$.

After finding R_X, we evaluate the probability of each event in R_X.

$\Rightarrow$ For $X = x$ $\{x = 0, 1, 2, 3\}$

$\Rightarrow$ $P(X=x)$ $\{\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}\}$

* Cumulative distribution function :-

Let X be a random variable with some probability distribution. then CDF of X is defined as,

$$F_X(x) = P(-\infty < X \leq x) \quad \forall x \in \mathbb{R}$$

$$= P(X \leq x) \quad \{\text{LHS of inequality is simply suppressed}\}$$

- Properties of CDF:-

(i) $0 \leq F_X(x) \leq 1 \quad \forall \ x \in \mathbb{R}$ (since CDF is also probability only)

(ii) $\lim_{x \to \infty} F_X(x) = F_X(\infty) = \underline{\underline{1}}$

$(F_X(\infty) = P(-\infty \leq X \leq \infty) \Rightarrow$ probability of full sample space)

(iii) Similarly, $F_X(-\infty) = \underline{\underline{0}}$.

(iv) $F_X(x)$ is non decreasing in x.

(v) $F_X(x)$ is right continuous at each x

$\Rightarrow (F_X(x+h) = F_X(x) \text{ as } h \to 0^+)$

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

$\Rightarrow P(X \geq x) = 1 - F_X(x)$

(vii) $P(a < X < b) = F_X(b) - F_X(a)$

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

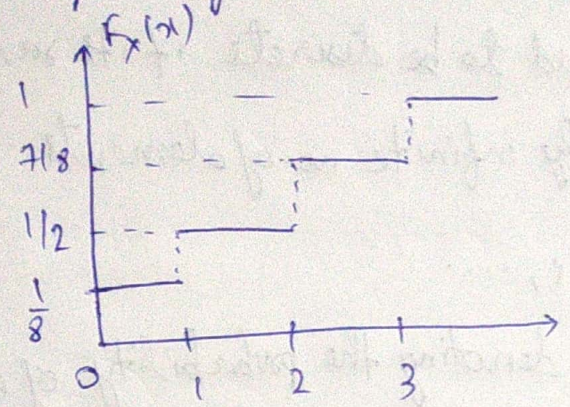
In general, if properties (ii), (iii), (iv) and (v) are satisfied for a given $F(x)$, then it must be the CDF of some random variable $X \Rightarrow F_X(x)$.

#

$X = x$	0	1	2	3
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$\Rightarrow F_X(x) = P(X \leq x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 1/2 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

$\Rightarrow$ Corresponding to this CDF, we have the graph,



⊛ The points where we have a jump are the points of positive probability, other points are of zero probability

$\Rightarrow \quad P(X=x) = F_X(x) - F_X(x^-)$

$\Rightarrow$ Here $\quad P(X=0) = F_X(0) - F_X(0^-) = \frac{1}{8} - 0 = \underline{\underline{\frac{1}{8}}}$

$P(X=1) = F_X(1) - F_X(1^-) = \frac{1}{2} - \frac{1}{8} = \underline{\underline{3/8}}$ ... and so on.

(where we don't have jumps $\Rightarrow P(X=x) = 0$).

$\Rightarrow$ We have found $P(X=x)$ from CDF $\Rightarrow$ reverse of what we did earlier.

Toss a fair die twice, X: Sum on upper faces

$\Rightarrow \quad R_X : \{2, 3, 4, \ldots 12\}$

Toss a coin until a head comes

X: No. of tosses to get that head

$R_X : \{1, 2, 3 \ldots \infty\}$

Now say $X > 4 \Rightarrow \{TTTTH, TTTTTH \ldots$ and so on $\}$

* Types of Random variables:-

(i) Discrete

(ii) Continuous

(iii) Mixed.

- Discrete Random variable:-

A random variable X is said to be discrete if its range space contains atmost countably infinite no. of elements like,

$$x_1, x_2, x_3, \dots x_i, \dots$$

Now let p_i be a number denoting the probability of event $X = x_i$, ie, $p_i = P(X = x_i)$, then the collection $\{p_i, i \in N\}$
$= P_X(x_i)$

is referred as the probability mass function (PMF) of X, provided that

(i) $P_X(x_i) \geq 0 \;\; \forall \; x_i \in R_X$

(ii) $\sum_{x_i \in R_X} P_X(x_i) = 1.$

$\Rightarrow$ $F_X(x)$ of a discrete random variable $= \sum_{x_i \leq x} P_X(x_i)$.

(we can check by applying definition of $F_X(x)$)

A shop has 10 products of a type of which 3 are defective. A buyer buys 2 products. Let X be the no. of defective products. then find PMF, CDF of X.

$\Rightarrow$ $R_X : \{0, 1, 2\}$

$\Rightarrow$ $P(X = x) \; \{ \frac{7}{10} \times \frac{6}{9}$

Scanned by CamScanner

* Mean and Variance of Random variables:-

Once we have the random variables and their probability distribution, we want a summary of information provided by these random variables and their probability distribution through mean, variance, median, quartile, etc.

- Mathematical Expectation (Average value) of a RVX :-

Let X be a random variable with some probability distribution. Then expected value of X is defined as,

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx$$ (X is continuous with prob. dist. func. $f_X(x)$)

provided, $\int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty$

If X is discrete with prob. distribution $\underbrace{P_X(x)}_{\text{prob. mass function}}$ then,

$$E(X) = \sum_{x_i \in R_X} x_i P_X(x_i)$$ provided, $\sum_{x_i \in R_X} |x_i| P_X(x_i) < \infty$

- Mean of a RVX may not always exist.

- In general, $E(g(X)) \xrightarrow{\text{if continuous}} \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

if discrete ($y(w) = f(x)$)

$$\sum_{x_i \in R_X} g(x_i) P_X(x_i)$$

- Mean is a measure of central tendency.

An example where mean may not exist :-

We take the case of Cauchy distribution for X (continuous).

For a Cauchy distribution, the $f_X(x)$ always $= \frac{\beta}{\pi}\left(\frac{1}{\beta^2+(x-\alpha)^2}\right)$

where $-\infty < x < \infty$
$-\infty < \alpha < \infty$
$0 < \beta < \infty$

$$\Rightarrow f_X(x) = \frac{1}{\pi}\left(\frac{1}{1+x^2}\right) \quad \left.\begin{matrix} -\infty < x < \infty \\ \beta = 1, \alpha = 0 \end{matrix}\right\} \text{Standard form of Cauchy distribution.}$$

$$\Rightarrow E(X) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x}{1+x^2}\,dx \qquad 1+x^2 = t \Rightarrow 2x\,dx = dt$$

$$\Rightarrow E(X) = \frac{1}{2\pi}\int \frac{dt}{t} = \frac{1}{2\pi}\int \frac{dt}{t} = \frac{1}{2\pi}\log t$$

$$= \frac{1}{2\pi}\log(1+x^2)\Big|_{-\infty}^{\infty} \Rightarrow \text{Non convergent.} \Rightarrow E(x) \text{ does not exist.}$$

$f_X(x) = \begin{cases} \frac{1}{x^2} & x \geq 1 \\ 0 & \text{o.w} \end{cases}$ Here also mean does not exist.

$f_X(x) = 1 \quad 0 < x < 1$
$\quad 0 \quad \text{o.w}$

$$E(X) = \int_0^1 x\,f_X(x)\,dx = \int_0^1 x\,dx = \underline{\underline{\frac{1}{2}}}.$$

$f_X(x) = \frac{1}{2}e^{-x/2} \quad x > 0$

$$\Rightarrow E(X) = \frac{1}{2}\int_0^{\infty} x e^{-x/2}\,dx = \underline{2}.$$

An example for discrete case:-

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$\Rightarrow E(X) = \sum_{x_i \in R_X} x \cdot p_X(x)$

$= 0 + \frac{3}{8} + \frac{3}{4} + \frac{3}{8} = \underline{\underline{\frac{3}{2}}}$. $\Rightarrow$ On an average, 1.5 heads will occur when a coin is tossed 3 times.

- Variance:-

Let X be a random variable with some probability distribution.

$\Rightarrow V(X) = E[(X - E(X))^2] \rightarrow$ average squared distance b/w data and the mean.

* Some properties of E:-
  - E is always linear
  - $E(ax+b) = aE(x) + b$

  $\Rightarrow$ when $a=0 \Rightarrow \underline{\underline{E(b) = b}} \Rightarrow$ Expected value of a constant = the constant.

$\Rightarrow V(x) = E[x^2 - 2xE(x) + (E(x))^2]$

$= E(x^2) - 2E(x) \cdot E(E(x)) + E((E(x))^2)$

$= E(x^2) - 2E(x) \cdot E(x) + (E(x))^2$

(since $E(x)$ is a constant)

$\Rightarrow \boxed{V(x) = E(x^2) - (E(x))^2} \Rightarrow V(x) \geq 0$ (since it is also an expectation only)

$$E(X^2) \xrightarrow{\text{discrete}} \sum_{x_i \in R_X} x_i^2 \, P_X(x_i)$$

continuous

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

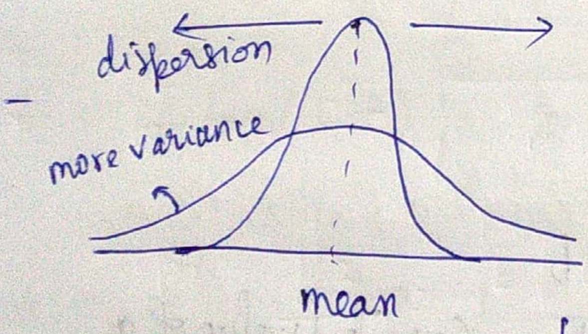
$f(x) = 1, \quad 0 < x < 1$

$\Rightarrow \quad E(x) = 1/2$

$\Rightarrow \quad E(x^2) = \int_0^1 x^2 dx = 1/3 \Rightarrow V(x) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$ (>0)

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

$\Rightarrow$ Only when $X = E(X) \Rightarrow$ for a constant x.

- 

Variance measures the dispersion. $\Rightarrow$ For the same centraltendency (mean) we can have different variance.

Mixed type example:-

$P_X(0) = 3/4$

$f_X(x) = \frac{1}{4} \quad 0 < x < 1$

$$E(X) = 0 \cdot P_X(0) + \int_0^1 x \cdot \frac{1}{4} dx$$

- We know,

$$E(ax+b) = aE(x) + b \quad (a, b) \text{ are given constants}$$

Now, $\boxed{V(ax+b) = a^2 V(x)}$ ⇒ Variance is not linear, unlike expectation.

Proof:- $V(ax+b) = E[ax+b - E(ax+b)]^2$

$$= E[ax+b - [aE(x)+b]^2]$$

$$= E[ax+b - a^2E^2(x) - b^2 - 2abE(x)]$$

Also, $\boxed{V(x+b) = V(x)}$ Shift in the data by a fixed constant has no effect on variance.

and $\boxed{V(ax) = a^2 V(x)}$

* A RV X is said to have a symmetric probability distn about a point α if $P(X \geq \alpha + x) = P(X \leq \alpha - x) \ \forall x \in \mathbb{R}$

$$\Rightarrow F_X(\alpha - x) = 1 - F_X(\alpha + x) + P(X = \alpha + x)$$

If X is continuous, $F_X(\alpha - x) = 1 - F_X(\alpha + x)$ CDF

$$\Rightarrow f_X(\alpha - x) = f_X(\alpha + x) \quad \left(\text{using } \frac{dF_X}{dx} = f_X\right)$$ PDF

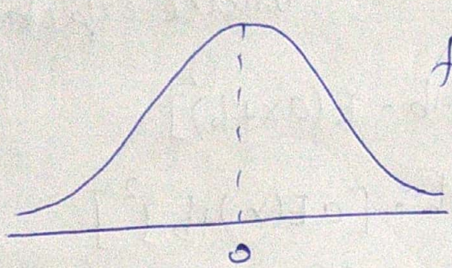
⇒ We get a simplified condition for continuous case.

(If $\alpha = 0$, $f_X(-x) = f_X(x) \ \forall x \in \mathbb{R}$)

$f_X(x) = \frac{\beta}{\pi} \frac{1}{\beta^2 + (x-a)^2}$

$\Rightarrow f_X(\alpha - x) = f_X(\alpha + x)$ about the point $\alpha = 0$

$\Rightarrow$



$f_X(x) = \frac{1}{\pi}\left(\frac{1}{1+x^2}\right) \quad -\infty < x < \infty$

Discrete example:-

$P_X(-1) = \frac{1}{4} = P_X(1) \quad , \quad P_X(0) = \frac{1}{2}$

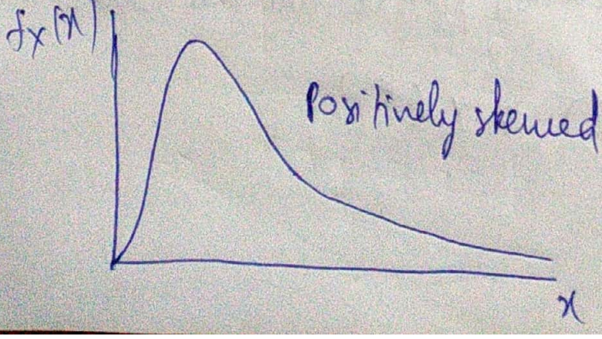
this distribution is symmetrical about 0. Here we have to check using $P(X \geq x) = P(X \leq -x) \; \forall x \in \mathbb{R}$.

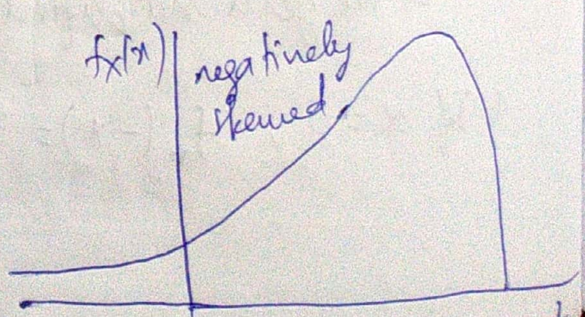
* Coefficient of Skewness:-

$$\beta_1 = \frac{\mu_3}{\sigma^3} = \frac{E[(X - E(X))^3]}{\sigma^3}$$

σ: standard deviation $= \sqrt{\text{variance}}$

$$= \begin{cases} 0 & \text{symmetric} \\ > 0 & \text{positively skewed} \\ < 0 & \text{negatively skewed} \end{cases}$$





- Positively skewed => Highly likely that the feature of probability we are looking for occurs in the beginning of the experiment itself.

- Negatively skewed :- Our feature of probability occurs near the end of the experiment (It is highly likely so).

- Symmetric :- Highly likely that it occurs somewhere in the middle of experiment.

* Quantile :-

Let X be a RV with some probability distribution.

A no. x_p is said to be the p^{th} quantile of this distribution if,

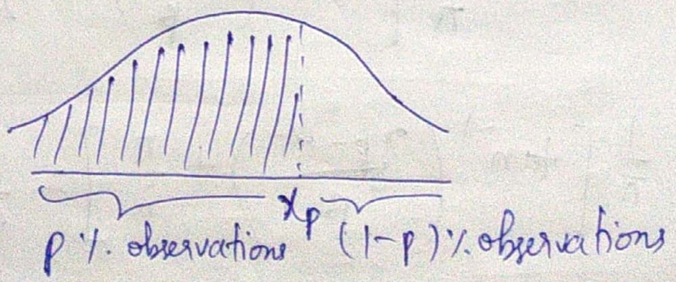
$$P(X \leq x_p) \geq p \qquad , \quad 0 < p < 1$$
$$P(X \geq x_p) \geq 1-p$$

$$\Rightarrow \quad p \leq F_X(x_p) \leq p + P(X = x_p)$$

Let X be continuous, then,

$$\boxed{F_X(x_p) = p}$$

$\Rightarrow$



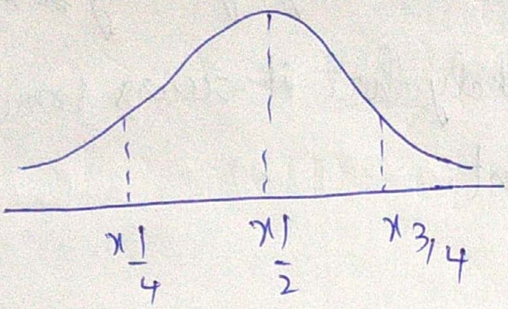
p^{th} quantile $x_p \Rightarrow p$% of the observations are lying behind x_p point on the real line, and $(1-p)$% above it.

Scanned by CamScanner

⇒ Using this logic, we can split any probability distribution into four equal sections,

$x_{\frac{1}{4}}, \; x_{\frac{1}{2}}, \; x_{3/4} \longrightarrow$ quartiles

1st quartile, 2nd, 3rd



⇒ We can also see that $x_{\frac{1}{2}}$ acts as the median for any probability distribution.

median → another measure of central tendancy.

\# $$f_X(x) = \frac{\beta}{\pi} \frac{1}{\beta^2 + (x-\alpha)^2} \qquad -\infty < x < \infty, \; \alpha \in R, \; \beta > 0$$

Find quartiles

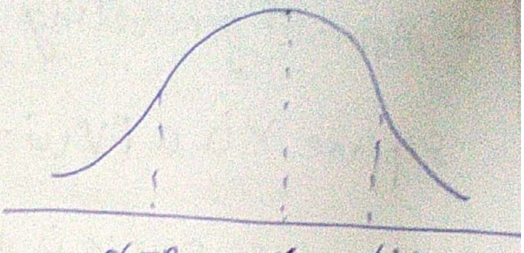
$$\Rightarrow F_X(x) = \frac{\beta}{\pi} \int_{-\infty}^{x} \frac{1}{\beta^2 + (t-\alpha)^2} \, dt$$

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

$$\Rightarrow \frac{1}{\pi} \left[\tan^{-1}\left(\frac{x_{\frac{1}{4}} - \alpha}{\beta}\right) + \frac{\pi}{2} \right] = \frac{1}{4}$$

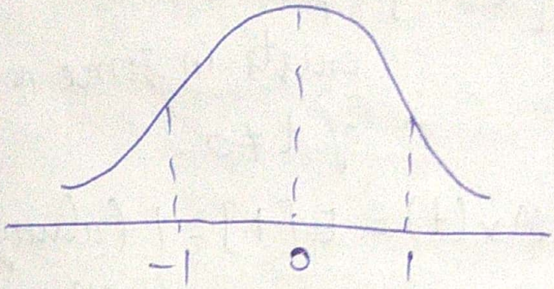
$$\Rightarrow x_{\frac{1}{4}} = \alpha - \beta$$ Similarly,

$x_{\frac{1}{2}} = \alpha$ and $x_{\frac{3}{4}} = \alpha + \beta$



$\Rightarrow$ for the cauchy distribution function, $\alpha = 0$, $\beta = 1$

$\Rightarrow$



Also, $\boxed{\text{Range of distribution} = x_{\frac{3}{4}} - x_{\frac{1}{4}}}$

Discrete case example:-

$P_X(-2) = \frac{3}{10}$ $\quad P_X(0) = \frac{1}{5}$ $\quad P_X(1) = \frac{1}{6}$ $\quad P_X(2) = 1/3$

Find median of this observation

$\Rightarrow$ We need $x_{\frac{1}{2}}$ such that,

$$P(X \geq x_{\frac{1}{2}}) \geq \frac{1}{2} \text{ and } P(X \leq x_{\frac{1}{2}}) \geq \frac{1}{2}$$

observe that
$\Rightarrow$ $P(X \geq 0) = \frac{1}{5} + \frac{1}{2} \geq \frac{1}{2}$

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

$P(X \leq 1) = \frac{2}{3}$

In fact, $[0,1]$

$\Rightarrow$ Any number between 0 and 1 can be verified as the median $(x_{\frac{1}{2}})$ $\Rightarrow$ this particular ex. has infinite medians.

* Moment generating function :- (MGF)

Suppose X is a RV with a given probability $dist^n$.

Then MGF of X is given by,

$M_X(t) = E[e^{tX}]$ provided this expectation exists in some neighbourhood of $t \neq 0$.

Because when $t=0$, $M_X(t) = E[1] = 1$ (always exists, so we move away from $t=0$).

$E[e^{tX}]$

discrete: $\sum_{x_i \in R_X} e^{tx_i} p_X(x_i)$

continuous: $\int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx$ (some parameter)

- $M_X(t)$ may not exist, since expectation does not always exist.

$f_X(x) = \frac{1}{\pi}\left(\frac{1}{1+x^2}\right) \quad -\infty < x < \infty$

Mean does not exist, similarly $M_X(t)$ also does not exist for this.

- $M_{aX+b}(t) = e^{bt} M_X(at)$ a,b → constants

$M_{2X+3}(t) = e^{3t} M_X(2t)$

- $\left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0} = E(X^n) \quad n=1,2,\ldots$

Discrete case :-

X	0	1	2	3
$P_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$\Rightarrow M_X(t) = 1 \cdot \frac{1}{8} + e^{t} \cdot \frac{3}{8} + e^{2t} \cdot \frac{3}{8} + \frac{e^{3t}}{8}$ { A function in t always }

$(t \neq 0)$

$\Rightarrow \frac{d}{dt} M_X(t) = \frac{3}{8}e^{t} + \frac{3}{4}e^{2t} + \frac{3}{8}e^{3t} \Big|_{t=0} = \frac{3}{2}$

$= E(X)$

$f_X(x) = 3e^{-3x}, \quad x > 0$

$\Rightarrow M_X(t) = \int_0^\infty e^{tx} \cdot 3e^{-3x} dx = 3\int_0^\infty e^{-(3-t)x} dx$

$= 3 \frac{e^{-(3-t)x}}{-(3-t)} \Big|_0^\infty = \frac{3}{3-t} \quad t < 3$

$\Rightarrow M'_X(t)\Big|_{t=0} = \frac{1}{3}$ and $3\int_0^\infty x e^{-3x} dx = \frac{1}{3} = E(X)$.

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

$= M''_X(t) - [M'_X(t)]^2$ at $t=0$.

- MGF uniquely determines a probability distribution if it exists.

* Probability Inequalities:-

(1) Marko Inequality :-

Let X be a nonnegative RV with finite mean. If $a>0$ be a given constant, then,

$$P(X>a) \leq \frac{E(X)}{a}$$

$E(X) \rightarrow$ mean of X

this is an estimate for the probability $P(X>a)$

Proof:-

$$\Rightarrow \quad E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_{0}^{\infty} x f_X(x)\,dx$$

($X \rightarrow$ non negative)

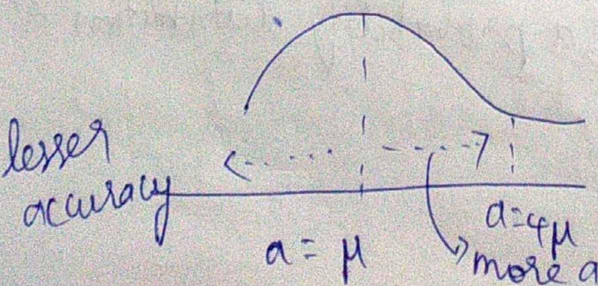
$$= \left(\int_0^a + \int_a^\infty\right) x f_X(x)\,dx$$

$$> a\int_a^\infty f_X(x)\,dx = aP(X>a)$$

Let $a = \frac{\mu}{2}$ where $\mu = E(X)$

$$\Rightarrow \quad P(X > \mu/2) \leq \frac{2\mu}{\mu}$$

$$\Rightarrow \boxed{P(X > \mu/2) \leq 2} \Rightarrow \text{lesser accuracy.}$$



$\Rightarrow$ this inequality will be very accurate for tail probability.

$f_X(x) = 3 \cdot e^{-3x}$, $x > 0$

$$\Rightarrow P(X > 2) = \int_2^{\infty} 3e^{-3x}\,dx = e^{-6}$$

$= 0.0\ldots$

Now lets estimate using marko inequality $\Rightarrow$

$E(X)$ can be calculated using $f_X(x) = \frac{1}{3}$

$$\Rightarrow P(X > 2) \leq \frac{1}{3(2)} \Rightarrow P(X > 2) \approx \frac{1}{6}$$ (not a great approx.)

(2) Tchebyshev Inequality:-

Let X be a RV whose mean and variance exist. If $k > 0$ is a given constant, then,

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

$\sigma^2 \rightarrow$ variance of X

$\mu \rightarrow$ mean of X

$(\Rightarrow P(a < X < b))$

Proof:- Let X be a continuous RV with pdf $f_X(x)$

$$\Rightarrow \sigma^2 = E(X - \mu)^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x)\,dx$$

$$= \int_{|x-\mu| \leq k} (x - \mu)^2 f_X(x)\,dx + \int_{|x-\mu| \geq k} (x - \mu)^2 f_X(x)\,dx$$

$$\geq \int_{|x-\mu| \geq k} (x - \mu)^2 f_X(x)\,dx \geq \underbrace{P(|X - \mu| \geq k)\,k^2}_{k^2 \int f_X(x)\,dx}$$

$k^2 \downarrow$ max. value of $(x - \mu)^2$ on the given range.

$\Rightarrow \quad P(|X-\mu| \geq k) \leq \dfrac{\sigma^2}{k^2}$ Thus proved.

\# $p_X(0) = 1 - \dfrac{1}{k^2}$, $k > 1$

$p_X(\pm 1) = \dfrac{1}{2k^2}$ Find an upper bound for the probability $P(|X| \geq 1)$ using Tchebyshev inequality

$\Rightarrow \quad P(|X-\mu| \geq k) \leq \dfrac{\sigma^2}{k^2}$

$\Rightarrow$ Here $E(X) = 0\left(1-\dfrac{1}{k^2}\right) + \dfrac{1}{2k^2} - \dfrac{1}{2k^2} = \underline{\underline{0}}$

$E(X^2) = 0\left(1-\dfrac{1}{k^2}\right) + \dfrac{1}{2k^2} + \dfrac{1}{2k^2} = \dfrac{1}{k^2}$

$\Rightarrow V(X) = E(X^2) - (E(X))^2 = \dfrac{1}{k^2}$

$\Rightarrow \sigma^2 = \dfrac{1}{k^2} \Rightarrow P(|X| \geq 1) \leq \dfrac{\frac{1}{k^2}}{1}$ this k is different from k in formula

$\Rightarrow$ An estimate for upper bound $= \underline{\underline{\dfrac{1}{k^2}}}$

and actual computation of $P(|X| \geq 1)$ also gives $= \underline{\underline{\dfrac{1}{k^2}}}$

$\Rightarrow$ Here we see that the estimate and actual value are same.

\# $f_X(x) = \dfrac{1}{2\sqrt{3}}$ $-\sqrt{3} < x < \sqrt{3}$ Find an U.B for $P(|X| \geq 3/2)$

compute exact result also.

* Some known discrete probability distributions:-

- Discrete uniform
- Binomial
- Poisson
- Geometric
- Negative Binomial
- Hyper Geometric.

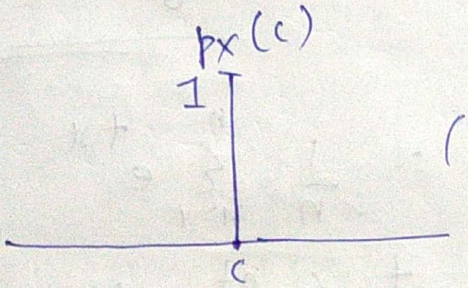
- Degenerate probability distribution:-

$P(X=c)=1 \Rightarrow$ entire probability mass is concentrated at a single point.

$\Rightarrow \quad E(X)=c$

$V(X)=E(X^2)-(E(X))^2=c^2-c^2=\underline{\underline{0}}.$

$M_X(t)=\underline{\underline{e^{ct}}}$ and $\underset{(CDF)}{F_X(x)}=\begin{cases} 0 & x<c \\ 1 & x\geq c \end{cases}$



(probability v/s random variable)

- Discrete uniform distribution:-

A RV X is said to have a discrete uniform $dist^n$ if its PMF is given by,

$\underset{(RV)}{P_X(x)}=\dfrac{1}{n}, \quad x=1,2,\dots n$ ($n \rightarrow$ parameter)

$\Rightarrow$ All outcomes are assigned equal probability

$\Rightarrow$ no outcome is more likely than any other.

$X \sim DU[1, n]$ is a notation to represent the discrete uniform distribution

$\underset{\text{discrete}}{D}\ \underset{\text{uniform}}{U}$

=) A given n describe all probabilistic information about the given distribution.

$$\sum_{x=1}^{n} p_X(x) = 1 \Rightarrow \text{It is a } \underline{\text{proper}} \text{ distribution.}$$

$$E(X) = \sum_{x=1}^{n} x\, p_X(x) = \sum_{x=1}^{n} \frac{x}{n} = \frac{1}{n}\sum_{x=1}^{n} x$$

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

$$E(X^2) = \sum_{x=1}^{n} x^2 p_X(x) = \frac{1}{n}\sum_{x=1}^{n} x^2 = \frac{n(n+1)(2n+1)}{n\ 6}$$

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

$$\Rightarrow V(X) = E(X^2) - (E(X))^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$$

$$M_X(t) = \sum_{x=1}^{n} e^{tx} p_X(x) = \frac{1}{n}\sum_{x=1}^{n} e^{tx}$$

$$\| $$

$$E(e^{tx})$$

$$= \begin{cases} \frac{e^t}{n}\left(\frac{e^{nt}-1}{e^t-1}\right), & t \neq 0 \\ 1 & t = 0 \end{cases}$$

and $F_X(x)$ (CDF) $= \begin{cases} 0 & x < 1 \\ 1/n & 1 \le x < 2 \\ 2/n & 2 \le x < 3 \\ \vdots & \vdots \\ 1 & x \geq n \end{cases}$

- Bernoullian Experiment :-

(1) Each trial results in two mutually disjoint and exhaustive outcomes known as Success and Failure.

(2) Trials are conducted independant of each other.

(3) Probability of getting a success at each trial remains fixed, lets call it p.

(4)* For binomial $dist^n$ no. of trials is finite.

· An example (application) of Discrete uniform :-

There are 100 slips of paper in a hat each of which has one of numbers 1, 2, ... 100 written on it. Five such slips are drawn at random one at a time (with replacement). Find prob. dist. of the value of jth draw ($1 \le j \le 5$). Find the prob. that number 100 is drawn atleast once.

Here $P(X=j)$ $(1 \le j \le 5) = \frac{1}{100}$ (prob. of drawing any slip $= \frac{1}{100}$)

$\Rightarrow$ $X \sim DU[1, 100]$

- Bernoulli distribution :-

conduct a Bernoullian expt. once.

X : no. of success in a single Bernoulli trial

R_X : 0, 1

0 → trial is failure

1 → the single trial resulted in a success

$\Rightarrow \quad P(X=1) = p$ (let)

then $P(X=0) = 1-p$ $\Rightarrow$ this can be written in a modified form as,

$$P_X(x) = p^x (1-p)^{1-x} \qquad \boxed{\begin{array}{l} x=0,1 \\ 0<p<1 \end{array}}$$

$$\Rightarrow \boxed{P_X(x) = p^x q^{1-x}} \quad \text{(Probability mass function)}$$

$$\Rightarrow E(X) = \sum_{x=0}^{1} x\, P_X(x) = 0\, P_X(0) + 1\, P_X(1)$$
$$= P_X(1) = \underline{\underline{p}}$$

$$E(X^2) = \sum_{x=0}^{1} x^2 P_X(x) = 0\, P_X(0) + 1\, P_X(1) = \underline{\underline{p}}$$

$$\Rightarrow V(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p)$$
$$= \underline{\underline{pq}}$$

$$M_X(t) = E(e^{tX})$$
$$= \sum_{x=0}^{1} e^{tx} P_X(x) = P_X(0) + e^t P_X(1)$$
$$= \underline{\underline{q + pe^t}}$$

(A MGF completely specifies information about the dist^n as it is unique to a dist^n)

$\Rightarrow \quad X \sim Ber(p) \rightarrow$ notation

we automatically know q.

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

- Binomial dist^n:-

Conduct Bernoullian expt. n number of times

X: no. of success in these n trials

R_X: 0, 1, 2, ... n → all trials are success

no success in any trial

$\Rightarrow \quad P_X(x) = {}^nC_x \, p^x \, (1-p)^{n-x}$

x success

n-x failure

there are total ${}^nC_x$ ways in which x success and (n-x) failure can occur in a total of n trials

$\Rightarrow$ PMF → $\boxed{P_X(x) = {}^nC_x \, p^x q^{n-x}}$ $\quad x = 0, 1, 2, 3 \ldots n$

$0 < p < 1$

$p + q = 1$

(n and p are the parameters)

[$X \sim Bin(n, p)$]

$$\sum_{x=0}^{n} {}^nC_x \, p^x q^{n-x} = (p+q)^n = \underline{\underline{1}} \Rightarrow \text{It is a proper distribution.}$$

$$E(X) = \sum_{x=0}^{n} x \, {}^nC_x \, p^x q^{n-x} = \sum_{x=1}^{n} x \, {}^nC_x \, p^x q^{n-x}$$

(as for $x = 0$, Σ(partial) $= 0$)

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x q^{n-x} \quad \text{put } y = x-1$$

$$= np \sum_{y=0}^{n-1} \frac{n-1!}{y!\,(n-1-y)!} p^y q^{n-y-1}$$

$$= np\,(p+q)^{n-1} = \underline{\underline{np}}$$

Scanned by CamScanner

Next we compute variance

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

Here $E(X^2) = \sum x^2 p_X(x) = \sum x^2 ({}^nC_x) p^x q^{n-x}$

$\Rightarrow$ Here one x will get cancelled, but the other remains and thus makes it difficult to simplify it, hence

for discrete case, $E(X^2) = E[X(X-1) + X]$

$$= E[X(X-1)] + E(X)$$

(k^{th} factorial moment $= E[X(X-1)\cdot(X-2)\ldots(X-k+1)]$)

$$\Rightarrow V(X) = E[X(X-1)] + E(X) - [E(X)]^2$$

$$= E[X(X-1)] + np - n^2p^2$$

$$\Rightarrow E[X(X-1)] = \sum_{x=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x q^{n-x}$$

$(x-2=y)$

$$= \sum_{y=0}^{n-2} \frac{n!}{y!(n-2-y)!} p^{y+2} q^{n-2-y}$$

$$= n(n-1)p^2 \sum_{y=0}^{n-2} ({}^{n-2}C_y) p^y q^{n-2-y}$$

$$= n(n-1)p^2 (p+q)^{n-2} = \underline{n(n-1)p^2}.$$

$$\Rightarrow V(X) = n(n-1)p^2 + np - n^2p^2 = n^2p^2 - np^2 + np - n^2p^2$$

$$= np - np^2 = np(1-p)$$

$\Rightarrow V(X) = npq$

coefficient of skewness $(\beta_1) = \dfrac{E[(X-E(X))^3]}{\sigma^3}$

$= \dfrac{npq(1-2p)}{(npq)^{3/2}} \Rightarrow \beta_1 = \dfrac{1-2p}{\sqrt{npq}} = \begin{cases} 0 & p=\frac{1}{2} \\ >0 & p<\frac{1}{2} \\ <0 & p>\frac{1}{2} \end{cases}$

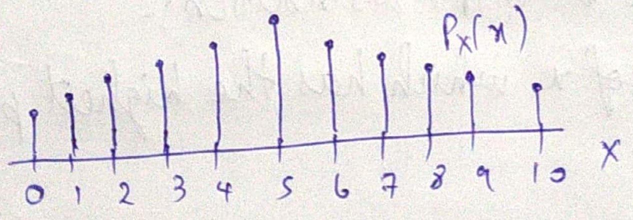
$\Rightarrow$ $p=\frac{1}{2} \Rightarrow$ Symmetric, $p<\frac{1}{2} \Rightarrow$ positively skewed and $p>\frac{1}{2} \Rightarrow$ negatively skewed.

$X \sim Bin(\underset{n}{10}, \underset{p}{1/2})$ $\Rightarrow$ Now we know all probabilistic information about the distribution

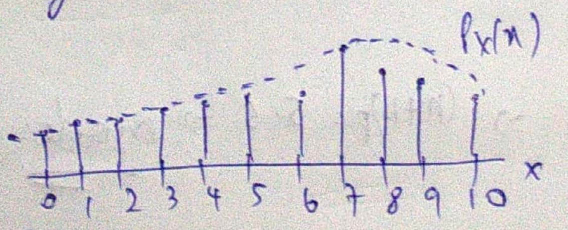
$\Rightarrow$ $P_X(x) = \left({}^{10}C_x\right)\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{n-x}$, $x = 0, 1, 2, 3 \ldots 10$

since $p=\frac{1}{2}$, we have symmetric distribution

$\Rightarrow$



$X \sim Bin(10, \frac{4}{5})$
negatively
(~~negatively~~ skewed)
~~positively~~



Since we have $\Rightarrow$ greater probabilities for higher x (x = no. of successes)

$\Rightarrow$ In this case, a ~~negatively~~ negatively skewed distribution gives us a higher probability of getting larger no. of success.

- MGF of a Bin(n, p) distribution :-

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{n} (e^t p)^x \, {}^nC_x \, q^{n-x}$$

$$\Rightarrow \quad \underline{M_X(t) = (q + pe^t)^n}$$

Let MGF of a RV X be given by

$$M_X(t) = \left[\frac{1}{5} + \frac{4}{5} e^t\right]^{10} \Rightarrow q = \frac{1}{5}, \; p = \frac{4}{5}, \; n = 10$$

$\Rightarrow$ Now we can get any information we want say $P_X(x \geq 3)$.

- $$\beta_2 = \text{coefficient of peakedness} = \frac{E[(X - E(X))^4]}{\sigma^2}$$

- Mode of a Bin(n, p) distribution :-

That value of x which has the highest probability of occuring

$\Rightarrow$ If $(n+1)p$ is not an integer, then $\underbrace{[(n+1)p]}^{\text{integral part}}$ is the mode

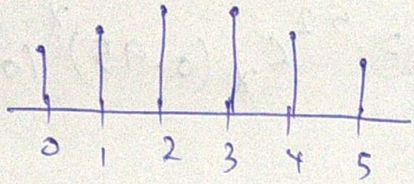
$\Rightarrow$ If $(n+1)p$ is an integer, then $(n+1)p$, $(n+1)p - 1$ are the two modes.

$\Rightarrow X \sim Bin(10, \frac{1}{2}) \Rightarrow (n+1)p = 5.5 \Rightarrow \underline{\text{mode} = 5}$

and $X \sim Bin(5, \frac{1}{2}) \Rightarrow (n+1)p = 3 \Rightarrow$ 2 and 3 are the modes

and we have $p = \frac{1}{2}$ (symmetric)

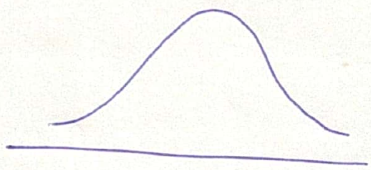
$\Rightarrow$



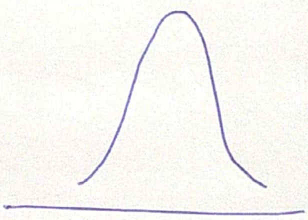
coefficient of peakedness (maximum height of our distribution)

$$\beta_2 = \frac{E((X - E(X))^4)}{\sigma^4} - 3$$

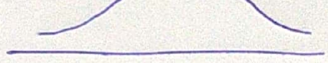
$= 0$ peak is like normal distribution



$> 0 \Rightarrow$ leptokurtic
(more than normal)



$< 0 \Rightarrow$ platikurtic
(less than normal)



An airline company knows that 5% of the people booking ticket do not turn up for a flight. So it sells 52 tickets for a 50 seat flight. What is the probability that every passenger who turned up for the flight will get a seat.

$\Rightarrow$ X : no. of passengers who turned up for the flight.

$\Rightarrow$ We have $Bin(52, 0.95)$ and we are interested

in $P_X(x \le 50) = 1 - P_X(x=51) - P_X(x=52)$

where $P_X(x) = {}^{52}C_x (0.95)^x (0.05)^{52-x}$