7/1/19

Inferential statistics :- Go beyond just merely reporting the data, conclusions are drawn from the data.

Experiment :- Conducting or observing something happen resulting in certain outcomes.

Deterministic expts :- If the cond^n's of expt. are fixed and then the outcome is known in advance then it is called deterministic expt.

Non deterministic expts. or Random expts.

If all the cond^n's are fixed even then the outcome cannot be predicted in advance then it is said to be random expt.

A random expt. can have several possible outcomes. A set of all possible outcomes of a random ~~expt~~ expt. is called a sample space. It is usually denoted by Ω or S.

Examples :- 1. Tossing of a coin.

$$\Omega = \{H, T\}$$

2. Tossing of a dice.

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

3. Drawing of a card

$$\Omega_1 = \{R, B\}.$$

$$\Omega_2 = \{Club, Diamond, Spade, Heart)$$

$$\Omega_3 = \{C_1, C_2, \ldots, C_{13}, S_1, \ldots, S_{13}, D_1, \ldots, D_{13}, H_1, \ldots, H_{13}\}$$

Birth of a child :- $\Omega_1 = \{B, G\}$.

We make our sample spaces according to our interests.

Amount of Rainfall in a region. $(50, 100)$ in cm.
Yield of wheat in a particular season $(1000, 5000)$.

100 mt sprint in an Olympics final

$\Omega_1 = \{A_1, A_2, \dots, A_8\}$
$\Omega_2 = (9, 10)$

Probability of an event :-

Event :- A event is a subset of the sample space.

Tossing of a dice : $E = \{$the number is even$\}$
$= \{2, 4, 6\}$

life of bulb is less than 10 days
$A = (0, 1/3)$

ϕ = Impossible event.
Ω = Sure event.

Union of events : $A \cup B$ — Occurence of A or B or both

$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^{n} A_i$
— Occurence of atleast one of A_i

Intersection of events :- $A \cap B$.
Both A and B occurs
Simultaneous occurence of A and B.

$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^{n} A_i$

simultaneous occurence of all A_i's.

If $\bigcup_{i=1}^{\infty} A_i = \Omega$, then A_i's said to be exhaustive event.

If $A \cap B = \phi \rightarrow A$ and B are mutually exclusive events.

$A^c \rightarrow$ Not happening of A.

$A - B \rightarrow$ happening of A but not of B.

$= A \cap B^c$

Mathematical or Classical Defn of Probability :-

Suppose a random expt. has N possible outcomes. Suppose these outcomes are equally likely & mutually exclusive. For an event E of these are M favourable outcomes, then the probability of E is defined as $P(E) = \frac{M}{N}$.

Limitations :-

1. N need not be finite.
2. The use of term 'equally likely' means that it is known that outcomes have equal 'chance' of appearing. Moreover, there must be prior knowledge of chances of outcomes.

Relative frequency or Statistical Defn of Probability :-

Suppose a random expt. is conducted a large number of times under identical condns. Suppose in n no. of trials of the expt. an event E occurs a_n times. Then, the probability of E is defined as $P(E) = \lim_{n \to \infty} \frac{a_n}{n}$

Probability are nos. attached to sets.

PAGE No.
DATE / /

Example.

H H T H H T . . . . .

$$\frac{a_n}{n} \rightarrow \frac{1}{1}, \frac{2}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots$$

$$\frac{a_n}{n} = \frac{2k}{3k} \quad n = 3k.$$

$$\frac{2k}{3k-1} \quad n = 3k-1.$$

$$\frac{2k-1}{3k-2} \quad n = 3k-2$$

$$k = 1, 2, \ldots$$

$$P(H) = \lim_{k \to \infty} \frac{a_n}{n} = \frac{2}{3}$$

Limitations of the defn :-

① Observations should be given to you. They observations may not be available.

② The expts. should be available to you.

Eg. You have to light all matchsticks for calculating the probability of defective matchsticks.

③ Order of a_n may be less than n.

Axiomatic defn of probability :-

[A. N. Kolmogorov (1933)]

Suppose we have a random expt. resulting in a sample space Ω.

A σ- algebra (σ-field) of subsets of Ω is a set $\mathcal{B}$ satisfying the following properties.

Sample Space → All sets of outcomes.

Event → Simply, subsets of sample spaces

and also $\phi \in \mathcal{B}$.

i) If $A_1, A_2, \ldots \in \mathcal{B}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$

ii) $A \in \mathcal{B} \Rightarrow A^c \in \mathcal{B}$

$(\Omega, \mathcal{B})$ is called a measureable space.

Defⁿ :- $(\Omega, \mathcal{B})$ is a probability space if we can define a funcⁿ $P: \mathcal{B} \to \mathbb{R}$ satisfying the following three axioms : matching each subsets of S $(\mathcal{B})$ with R.

A_1 : $P(E) \geq 0 \quad \forall \ E \in \mathcal{B}$

A_2 : $P(\Omega) = 1$.

A_3 : If $E_1, E_2, \ldots$, are pairwise disjoint events in $\mathcal{B}$, then.

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

$(\Omega, \mathcal{B}, P) \to$ Probability space.

Some properties of Probability functions :-

P1 : $P(\phi) = 0$

Proof:- let $E_1 = \Omega$ and $E_2 = E_3 = \ldots = \phi$

Then,

$$P\left(\bigcup_{i=0}^{\infty} E_i\right) = \sum_{i=0}^{\infty} P(E_i)$$

$$\Rightarrow P(\Omega) = P(\Omega) + P(\phi) + P(\phi) + \ldots$$

$$\Rightarrow 1 = 1 + P(\phi) + P(\phi) + \ldots$$

$$\Rightarrow P(\phi) + P(\phi) + P(\phi) + \ldots = 0$$

$$\boxed{\therefore P(\phi) = 0}$$

P2 : If $E_1, \ldots, E_n$ are pairwise disjoint event in $\mathcal{B}$, then

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

For finite sets :- $\mathcal{B} = \{$ All the subsets of S, including S itself $\}$

For ∞ sets :- $\mathcal{B} = \{$ All sets of the form $[a,b]$, $(a,b]$, $[a,$ $(a,$ $\forall\, a, b \in \mathbb{R}$.

Choose.

$$E_{n+1} = E_{n+2} = \dots = \phi.$$

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) + P(\phi) + P(\phi) + \dots$$

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

P_3 :- Probability func^n is monotone,

i.e. if $E \subset F$, then $P(E) \le P(F)$

$$F = E \cup (F - E)$$

$$P(F) = P(E) + \underbrace{P(F-E)}_{\geq 0}$$

$\therefore$ $P(F) \geq P(E)$

Also, $P(F-E) = P(F) - P(E)$

P_4: For any event E, we have
$$0 \le P(E) \le 1$$

Since, $\phi \subset E \subset \Omega$

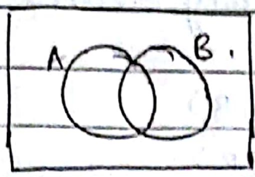
$\Rightarrow P(\phi) \le P(E) \le P(\Omega) \Rightarrow \underline{0 \le P(E) \le 1}$.

P_5: $P(E^c) = 1 - P(E)$.

Addition rule :-

For any two events $A, B \in \mathcal{B}$

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



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

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

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

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

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

$$+ \sum\sum\sum P(A_i \cap A_j \cap A_k) - \cdots$$

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

Theorem :- (Subadditivity prop. of Probability).

Let $A_1, A_2, A_3, \ldots, A_n \in \mathcal{B}$, then

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

Conditional Probability :-

Let $A, B \in \mathcal{B}$ with $P(B) > 0$. The cond^{n}al prob.^{n} of event A given that B has already occured is defined as:-

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

We prove that the conditional probability func^{n} $P(\cdot|B)$ is a valid probability f^{n}.

A_1: $P(A|B) \geq 0$ is always true.

A_2: $P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

A_3: Let $A_1, \ldots, A_n$ be pairwise disjoint events in $\mathcal{B}$.

$$P\left(\bigcup_{i=1}^{\infty} A_i | B\right) = \frac{P\left(\bigcup_{i=1}^{n} A_i \cap B\right)}{P(B)}$$

$$= \sum_{i=1}^{n} \frac{P(A_i \cap B)}{P(B)}.$$

$$= \sum_{i=1}^{n} P(A_i | B).$$

Multiplication rule :-

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

PAGE No.
DATE / /

General Multiplication Rule :- Let $A_1, A_2, \ldots, A_n$ be events in $\mathcal{B}$ with $P(\bigcap_{i=1}^{n} A_i) > 0$

Then,

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

Proof :- (By induction.)

Theorem of total probability :-

Let $A \in \mathcal{B}$ and $B_1, B_2, \ldots, B_n \in \mathcal{B}$. Assume B_i's are mutually exclusive and exhaustive events with $P(B_i) > 0$, $i = 1, \ldots, n$. Then, with $P(B_i) > 0$, $i = 1, \ldots, n$.

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

Proof :- $\bigcup_{i=1}^{n} B_i = \Omega$.

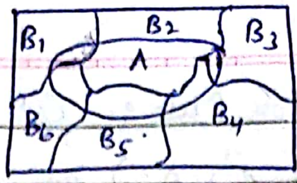
$$P(A) = P(A \cap \Omega) = P\left(A \cap \left(\bigcup_{i=1}^{n} B_i\right)\right)$$

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

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

Probability of the cause due to all possibilities.

<u>Baye's thm.</u>



Let $A \in \mathcal{B}_0$ with $P(A) > 0$ and $B_1, B_2, \ldots, B_n \in \mathcal{B}$. Assume $B_1, \ldots, B_n$ are mutually exclusive and exhaustive events. Then,

$$P(B_j|A) = \frac{P(A|B_j)\, P(B_j)}{\underbrace{\sum_{i=1}^{n} P(A|B_i)\, P(B_i)}_{\text{Total probability}}}$$

Proof :-

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

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

<u>Independence of Events</u> :-

If A doesn't depend on B.

$$P(A|B) = P(A)$$

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

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

So we define A and B to be independent events if
$P(A \cap B) = P(A)P(B)$.

Three events A, B, C are said to be mutually independent of
$P(A \cap B) = P(A)\,P(B)$, $P(B \cap C) = P(B)P(C)$,
$P(C \cap A) = P(C)P(A)$, $P(A \cap B \cap C) = P(A)\,P(B)\,P(C)$.

The total no. of cond^{n}s is $2^n - n - 1$. (for n sets)

Problem :-

Q.) Birthday problem :- Suppose there are n persons in a hall and none has bday on 29th Feb. (Let $n \le 365$). What is the probability that atleast two persons have same birthday.

→ More easy to calculate when no person has same bday.

$$P(A^c) = \frac{{}^{365}P_n}{(365)^n}$$

$A^c \longrightarrow$ no two persons have same b'day.

$$P(A) = 1 - \frac{{}^{365}P_n}{(365)^n} = 1 - \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\cdots\left(1 - \frac{n-1}{365}\right)$$

Q.) Six cards are drawn with replacement from an ordinary deck. What is the prob. that each of the four suits will be represented atleast once among the six cards?

52 cards.

A.) $\cancel{\frac{(3\times 13)^6}{52^6}} + \cancel{\frac{(2\times 13)^6}{(52)^6}} + \cancel{\frac{(1\times 13)^6}{(52)^6}}$.

$\cancel{A^c} = \left(\frac{3}{4}\right)^6 + \cancel{\left(\frac{2}{4}\right)^6} + \cancel{\left(\frac{1}{4}\right)^6}$ $\rightarrow \cancel{P}(A) = 1 \cancel{-} \cancel{P(A^c)}$

PAGE No.

DATE / /

$A \rightarrow$ All suits did appear.

$A^c \rightarrow$ Atleast one suit doesn't appear.

$$A^c \rightarrow \bigcup_{i=1}^{4} B_i$$

$B_1 \rightarrow$ spades do not appear.

$B_2 \rightarrow$ hearts do not appear.

$B_3 \rightarrow$ diamonds do not appear.

$B_4 \rightarrow$ clubs do not appear.

$$P(A^c) = \sum_{i=1}^{n} P(B_i) - \sum\sum P(B_i \cap B_j)$$

$$+ \sum\sum\sum P(B_i \cap B_j \cap \cdots \cap B_n)$$

$$= 4 \cdot \left(\frac{3}{4}\right)^6 - 6\left(\frac{1}{2}\right)^6 + 4\left(\frac{1}{4}\right)^6$$

$$= \frac{317}{512} \approx 0.62$$

$$P(A) = 1 - 0.62 = 0.38$$

Q.) If 4 married couples are arranged to be seated in a row, what is the prob. that no husband is seated next to his wife?

H_1 H_2 H_3 H_4.

W_1 W_2 W_3 W_4.

X X X X X X X X.

Scanned by CamScanner

$A \rightarrow$ When atleast one husband is sitting next to his wife.

$$E^c = \bigcup_{i=1}^{4} A_i \qquad A_i = i^{th} \text{ couple is together.}$$

$$P(A_i) = \frac{7! \times 2}{8!} = \frac{7 \times 2}{8 \times 7} = \frac{1}{4}. \quad i = 1, \ldots, 4.$$

$$P(A_i \cap A_j) = \frac{6! \times 2^2}{8!} = \frac{4}{8 \times 7} = \frac{1}{14}.$$

$$P(A_i \cap A_j \cap A_k) = \frac{5! \times 2^3}{8!}$$

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \frac{4! \times 2^4}{8!}$$

$$P\left(\bigcup_{i=1}^{4} A_i\right) = \frac{23}{35} \qquad P(E) = \frac{12}{35} \approx 0.34.$$

Q.> Three players A, B and C take turns in throwing a dice in order ABC, ABC, ... What is the probability that.

i> A is the second player to get a 6 for the first time,

ii> A is the last player to get a six for the first time?

$A \rightarrow$ Player getting 6 at i^{th} cycle. $P(\text{Getting } 6) = \frac{1}{6}$.

$P(\text{Not getting } 6) = 5/6$

A, B, C, A, B, C, ...

A gets to throw on $(3r+1)^{th}$ trial $r = 0, 1, 2, 3, \ldots$

Suppose he/she gets a six on the final trial.

In previous r trials he/she fails.

So this prob. $= \left(\frac{5}{6}\right)^r \cdot \frac{1}{6}$

B may get a six in r trials with prob.

$$1 - \left(\frac{5}{6}\right)^r$$

C may not get a six in r trials with prob.

$$\left(\frac{5}{6}\right)^r$$

So, the reqd. prob. is

$$2 \sum_{r=1}^{\infty} \left(\frac{5}{6}\right)^{2r} \left\{1 - \left(\frac{5}{6}\right)^r\right\} \cdot \frac{1}{6} = \frac{300}{1001} \approx 0.2997.$$

$\frac{~~395~~}{1001}$ next ans.

It is often imp. to allocate a number to the outcome.

a X

Random variables :- Random variable is basically a f^n from sample space to real nos. $X : \Omega \to R$.

Quite often in a random expt., we assign real values to the outcomes. Say, we define a real valued $func^n$.

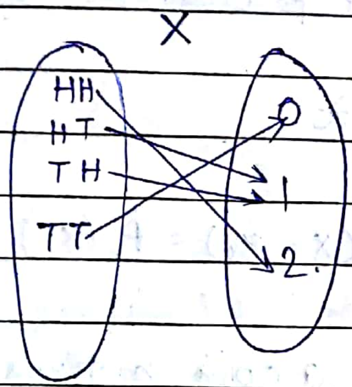
$$X : \Omega \to R.$$

i.e. X is said to be a random variable if it assign real values to the elements of the sample space.

Eg:- 1. Suppose n coins are tossed. Let X → the number of heads. Then X is the random variables.

It takes values 0, 1, 2, 3, ..., n.

Let n = 2



$X(HH) = 2$

$X(HT) = X(TH) = 1$

$X(TT) = 0$

2. Let us consider the life of a bulb (in hrs.)

X → the life. $\Omega = [0, \infty)$

Then X is a random variable & $X \in [0, \infty)$

3. Tossing of two dies → X_1 = the sum.

X_2 = the product.

→ Discrete Random Variable. → When Range (X) is countable.

→ Continuous Random Variable.

Scanned by CamScanner

PMF is defined for values in the range, it is convenient to extend PMF of X to all real nos.

If a random variable takes finite or countably infinite no. of values, it is called a discrete random variable.

If the range of a random variable is an interval, it is called a continuous random variable.

Probability Distribution :- of a Random Variable.

Let X be a discrete random variable taking values in set $\mathcal{X} = \{x_1, x_2, x_3, \ldots\}$

The probability distribution of X is defined by a func^n $p_X(x_i)$ (called probability mass func^n) satisfying the following properties:-

1. $p_X(x_i) \geq 0 \quad \forall\ x_i \in \mathcal{X}$

2. $\sum_{i=1}^{\infty} p_X(x_i) = 1$ 3. $P(X = x_i) = p_X(x_i)$

Q.) In the expt. of tossing a coin n times.

Suppose $P(\text{Head in a single toss}) = p$

$$P(X=0) = (1-p)^r$$
$$P(X=r) = {}^nC_r \, p^r (1-p)^{n-r}$$

$$p_X(r) = P(X=r) = {}^nC_r \, p^r (1-p)^{n-r}$$

$r = 0, 1, 2, \ldots$

PAGE No. DATE / /

Probability distribution of a continuous Random Variable.

Let X be a continuous random variable taking values on $\mathbb{R} = (-\infty, \infty)$. The prob. distribution of X is described by a func^n $f_X(x)$ (called probability density func^n) or pdf.

satisfying the following properties :-

1. $f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.
3. $P(a < x < b) = \int_a^b f_X(x)\, dx$

Eg :- Let X be the life of a bulb and it has been found that

$$f_X(x) = \begin{cases} 3e^{-3x} & , x > 0 \\ 0 & , x \leq 0. \end{cases}$$

Mixed Random Variable :-

Sometimes a random variable X may have +ve probabilities for some points and pdf over some intervals. Such a r.v is called a mixed variable

Eg :- Waiting time for chapati at lunchtime (in mins.) in the mess. Then $P(X=0) = \frac{1}{10}$

$$f_X(x) = \begin{cases} \frac{9}{10} \cdot \frac{1}{20} & 0 < x < 20 \\ 0 & , o/w. \end{cases}$$

Cumulative Distribution Funcⁿ :-

Let X be a r.v., we define.

$$F_X(x) = P(X \leq x) \quad , \quad x \in \mathbb{R}.$$

In case X is a discrete r.v. with values in $\mathcal{X} = \{x_1, x_2, \ldots\ldots\}$, then,

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

Conversely,

$$p_X(x_i) = P(X = x_i) = P(X \leq x_i) - P(X \leq x_{i-1})$$
$$= F_X(x_i) - F_X(x_{i-1})$$

In case X is a continuous r.v with pdf $f_X(x)$ then

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

Conversely,

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

Characterization properties of CDF :-

If $F_X(x)$ is a cdf of a r.v. Then,

1. $\lim_{x \to -\infty} F_X(x) = 0.$
2. $\lim_{x \to \infty} F_X(x) = 1.$
3. If $x_1 < x_2$ then $F_X(x_1) \leq F(x_2).$

4.) F_X is continuous from right at every point i.e. $\lim_{h \to 0^+} F_X(x+h) = F_X(x)$ $\forall x \in \mathbb{R}$.

Conversely, if a funcⁿ F satisfies the above four properties, then F is a cdf of some r.v X.

Example:- ① X = no. of heads in two tosses of a coin prob.(head) = $\frac{2}{3}$.

$X = 0, 1, 2$

$p_X(0) = \frac{1}{9}$, $p_X(1) = \frac{4}{9}$, $p_X(2) = \frac{4}{9}$

$$F_X(x) = 0 \qquad x < 0.$$
$$= 1/9 \qquad 0 \le x < 1.$$
$$= 5/9 \qquad 1 \le x < 2.$$
$$= 1 \qquad 2 \le x.$$

② $f(x) = \begin{cases} 3e^{-3x} & , x > 0 \\ 0 & , x \le 0 \end{cases}$.

$$F_X(x) = \int_0^x 3e^{-3t}\,dt = 1 - e^{-3x}$$

$$\therefore F_X(x) = \begin{cases} 0 & , x \le 0 \\ 1 - e^{-3x} & , x > 0 \end{cases}.$$

③ For Chapati problem :-

$$F_X(x) = 0 \quad , \quad x < 0$$
$$= \frac{1}{10} \quad , \quad x = 0$$
$$= \frac{1}{10} + \int_0^x \frac{9}{200}\,dt \quad , \quad 0 < x < 20$$
$$= 1 \quad , \quad x \ge 20$$

$$F_X(x) = 0, \quad x < 0$$

$$= \frac{1}{10} + \frac{9x}{200}, \quad 0 \le x < 20$$

(more informative??) $= 1. \quad x \geq 20$

Mathematical expectation of a R.V :-

(weighted mean of random variables)

Let X be a discrete random variable with pmf $p_X(x_i)$, $x_i \in \mathcal{X} = \{x_1, x_2, \dots\}$.

Expected value (Avg.) or mathematical expectation of X is defined as,

$$E(X) = \sum_{i=1}^{\infty} x_i \, p_X(x_i)$$

provided the series on the right is absolutely convergent.

If X is continuous r.v. with pdf $f_X(x)$ we define

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

provided the integral on the right is absolutely convergent.

E (time of getting chapatti)

$$= 0 \cdot \frac{1}{10} + \frac{9}{200} \int_0^{20} x \, dx.$$

$$= 9 \text{ mins.}$$

If all random variables

22/11/15

Example :- (Kulisyaly)

pdf.
pmf.
cdf.

Suppose a box has 4 bulbs out of which one is defective. Bulbs are tested one by one without replacement to identify the defective bulb.

A.) $X \rightarrow$ no. of testings reqd.

$P(X=1) \rightarrow 1/4$.

$P(X=2) \rightarrow \frac{3}{4} \times \frac{1}{3} = \frac{1}{4}$.

$P(X=3) \rightarrow \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2}$.

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

$\therefore E(X) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{2} = \frac{1}{4} + 2 = \frac{9}{4}$.

$F(x) = \frac{2}{\pi} \tan^{-1} x \quad , \quad -\infty < x < \infty$.

Ex :- (Convert it to CDF).

Generalization of the concept of expectation.

Let X be a r.v. with pmf / pdf and g be a real valued funcn. $g: R \rightarrow R$.

Then,

$Y = g(X)$ is a r.v.

$$E(Y) = E(g(X)) = \begin{cases} \sum_{x_i \in \mathcal{X}} g(x_i) \cdot p(x_i) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} g(x) f_X(x) \, dx & \text{if } X \text{ is continuous.} \end{cases}$$

PAGE No.
DATE / /

Mean or expected value of a random variable X is of special importance in statistics because it describes where the p.d is centered.

→ But the mean doesn't convey any info about the shape of the curve.

We must have the condⁿ of absolute convergence satisfied.

$g(X) = aX + b$ where a, b are constants.

$$E(aX+b) = \sum_{x_i \in \mathcal{X}} (ax_i + b) \cdot p_X(x_i)$$

$$= a \sum_{x_i \in \mathcal{X}} x_i \, p(x_i) + b \sum_{x_i \in \mathcal{X}} p_X(x_i)$$

$$= a\,E(X) + b$$

So, $\boxed{\text{expectation is a linear func}^n.}$

Moments :- $\mu'_k = E(X^k)$ $k = 1, 2, \ldots$

are called non central moments of X.

In particular, $\mu'_1 = E(X)$ is the avg. or mean values of X.

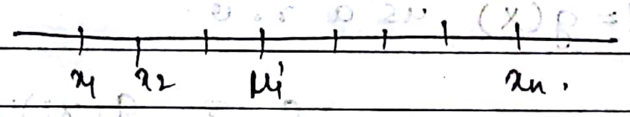
$\mu_k = E(x - \mu'_1)^k$. $k = 1, 2, \ldots$

are called central moments or moments about the mean of X.

?? $\mu_1 = E(x - \mu'_1) = E(X) - \mu'_1 = \mu'_1 - \mu'_1 = 0$.

So, the first moment is always zero.

$\mu_2 = E(X - \mu'_1)^2$ is called the variance of X.



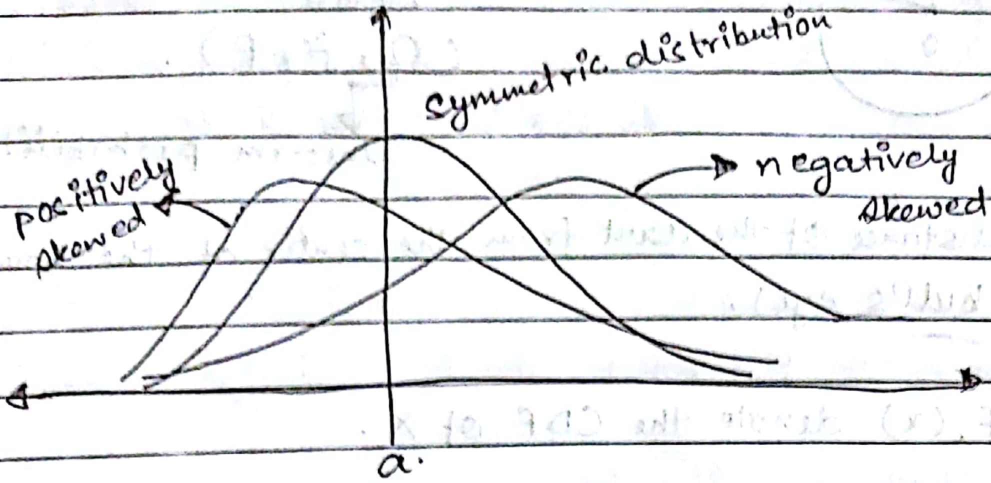
$= Var(X)$

Probability is just the relative frequency in long run.

(Measure of variability)

Variance $\rightarrow \sigma^2$

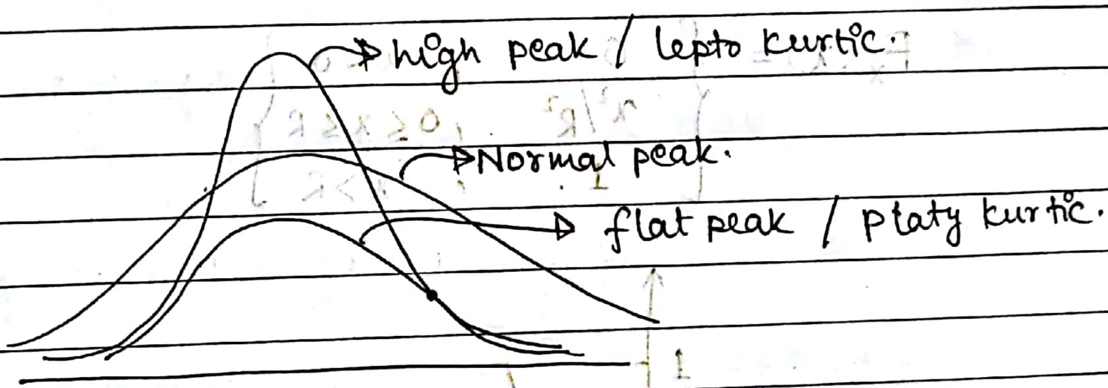
$SD(X) = \sqrt{VAR(X)}$ = Standard Deviation of X.



We define:- the sign of μ_3 is an indicator of the skewness of a distribution.

$$\boxed{\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}} \rightarrow \text{measure of skewness.}$$

β_1 has no units.



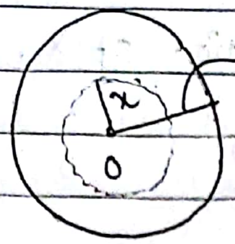
$$\beta_2 = \left(\frac{\mu_4}{\mu_2^2} - 3\right) \rightarrow \text{Measure of kurtosis or peakedness.}$$

The variance of a random variable,

$$\sigma^2 = \mu_2 = E(X-\mu_1')^2 = \underline{\underline{E(X^2) - \mu_1'^2}}.$$

28/1/19

Ex:-



Expt:- Throw dart on this circular board.

$(\Omega, \mathcal{F}, P)$.
$\downarrow$
Uniform probability space.

X: distance of the dart from the center of the board (bull's eye).

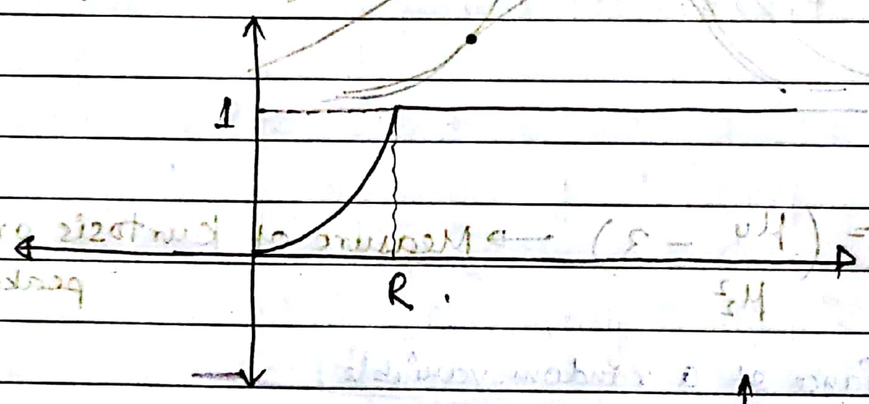
Let $F_X(x)$ denote the CDF of X.

$$F_X(x) = \text{Prob}(X \leq x)$$

$$= \frac{\pi x^2}{\pi R^2} \longrightarrow \text{Because of uniformity principle.}$$

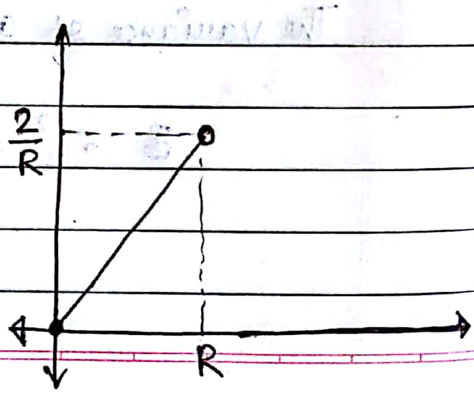
$$\boxed{F_X(x) = \frac{x^2}{R^2}}$$

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



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

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



PAGE No.
DATE / /

$$P(a \le x \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$$

For continuous variable, pdf does not tells us probability at a particular point.
For discrete variable, pmf tells us the probability at a particular point.

Let X be a continuous random variable with CDF $F_X(x)$ and $f_X(x) = \frac{d}{dx}F_X(x)$, is its probability density func^n (pdf).

range of a random variable.

Then the range of X is the $R_X \subseteq R$ such that $\forall x \in R_X, f_X(x) > 0$

Def^n:- A density func^n or pdf $f_X(x)$ is an non-ve func^n such that $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$.

Then, obviously, $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$. satisfies all properties of CDF.

$$P(a \le x \le b) = \int_a^b f_X(x)\,dx = F_X(b) - F_X(a).$$

Examples :-

1) For what values of k, $f(x) = kx^3$, $0 \le x \le 1$.
$= 0$ o/w.

is a pdf.

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

$$\int_0^1 kx^3\,dx = \left.\frac{kx^4}{4}\right|_0^1 = 1. \Rightarrow k = 4. \quad f_X(x) = \begin{cases} 4x^3, & x \in [0 \\ 0, & o/w \end{cases}$$

$$CDF = \begin{cases} x^4, & x \in [0,1] \\ 1, & o/w \end{cases}$$

Scanned by CamScanner

CDF value at x gives the $P(X \le x)$. (Imp.)

Change of variable formula :-

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

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

$$G_Y(y) = P(Y \le y) = P(X^2 \le y)$$ For any $y \in \mathbb{R}$.

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

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

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

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

$$= \frac{1}{2\sqrt{y}}\left[F_X'(\sqrt{y}) + F_X'(-\sqrt{y})\right]$$

$$g(y) = \frac{1}{2\sqrt{y}}\left[\underbrace{f_X}_{\text{Probability density } f^n}(\sqrt{y}) + f_X(-\sqrt{y})\right], \quad y > 0$$

Ex:

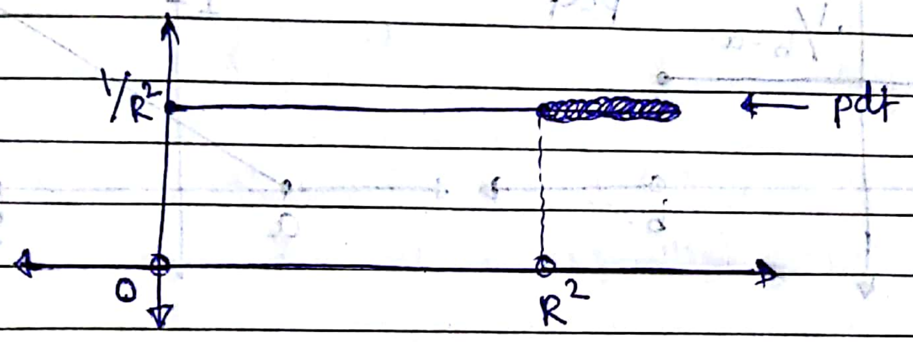
$$f_X(x) = \begin{cases} 2x/R^2 & , 0 \le x \le R. \\ 0 & , o/w. \end{cases}$$

Find the density of $Y = X^2$.

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

$$g(y) = \begin{cases} \frac{1}{2\sqrt{y}} \times \frac{2\sqrt{y}}{R^2} & , 0 \le y \le R^2 \\ 0 & , o/w \end{cases}$$

$$\rightarrow g(y) = \begin{cases} 1/R^2 & , 0 \le y \le R^2 \\ 0 & , o/w \end{cases}$$ Uniform continuous density.



Continuous uniform Random Variable :-

Let X be a continuous random variable with pdf defined as,

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

for any two $a, b \in \mathbb{R}$ such that $a \le b$. Then X is said to follow uniform density with parameter a & b.

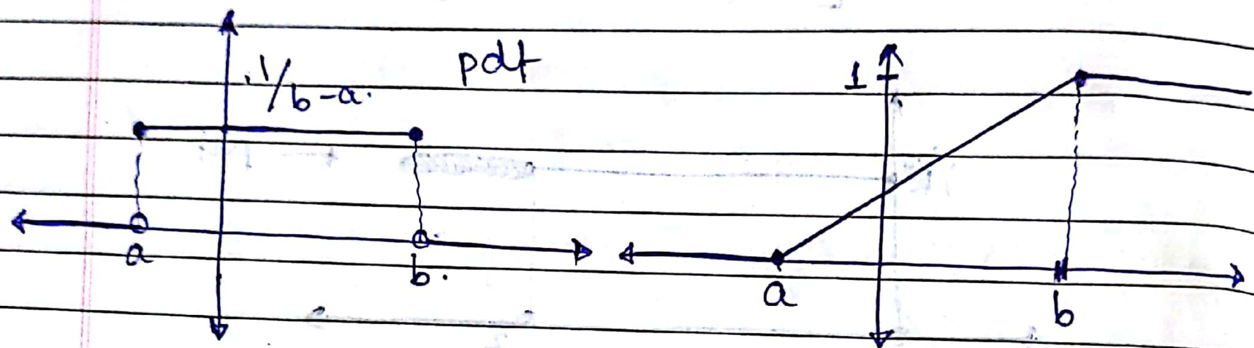
$$X \sim U(a, b).$$

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

$$= \int_{a}^{x} \frac{1}{b-a}\,.dx. \quad , \quad x \leq b$$

$$F_X(x). = \frac{x-a.}{b-a} \quad , x < b$$

$$F_X(x) = \begin{cases} 0 & , \; x < a \\ \frac{x-a.}{b-a.} & , a \leq x < b. \\ 1. & , \; x \geq b \end{cases}.$$



Ex:- Let $X \sim U(0,1)$

$$f_X(x) = \begin{cases} 1 & , \; 0 \leq x \leq 1 \\ 0 & , \; o/w \end{cases}$$

$$F_X(x) = \begin{cases} 0 & , \; x \leq 0 \\ x & , \; 0 \leq x < 1 \\ 1 & , \; x \geq 1 \end{cases}$$

Consider the transformation :-

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

A.} Let $G_Y(y)$ be the CDF of Y and $g(y)$ be the pdf of Y.

$$G_Y(y) = P(Y \le y).$$
$$= P\left(-\frac{1}{\lambda}\log(1-X) \le y\right).$$
$$= P\left((1-X) \le e^{-\lambda y}\right).$$
$$G_Y(y) = P\left(X \le 1-e^{-\lambda y}\right) = F_X(1-e^{-\lambda y}).$$

as, $1-e^{-\lambda y} \le 1$.

$$G_Y(y) = \begin{cases} 1-e^{-\lambda y}, & y>0 \\ 0, & y<0 \end{cases}.$$

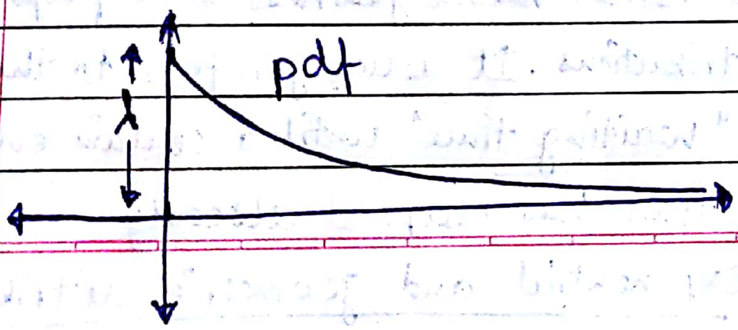
$$g_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y>0 \\ 0, & y \le 0 \end{cases}.$$

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

Ex :- Let $X \sim \exp(\lambda)$.
$\sim$ → follows.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x>0 \\ 0, & \text{o/w} \end{cases}.$$

$$F_X(x) = 0, \quad x<0$$
$$= 1-e^{-\lambda x}, \quad x \ge 0.$$



Scanned by CamScanner

For any real nos., $x, y > 0$.

$$P(X > x+y) = 1 - F_X(x+y)$$

$$= 1 - (1 - e^{-\lambda(x+y)})$$

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

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

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

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

→ Memory less property of exponential density.

Thm:- Let ϕ be a differentiable f^n which is strictly increasing or strictly decreasing on an interval I. Let $\phi(I)$ denote the range of ϕ and ϕ^{-1} be the inverse of ϕ on I. Let X be a continuous random variable having density $f_X(x)$ such that $f_X(x) \neq 0$ on I. Then $Y = \phi(X)$ whose density is given by $g(y) = f_X(\phi^{-1}(y)) \cdot \left|\frac{d}{dy}\phi^{-1}(y)\right|$

$y \in \phi(I)$.

$f_X(x) \neq 0$ on I.

means, I is in the range of X.

Probabilities wouldn't be influenced by the history of the expt.

In probability and statistics, memorylessness is a property of certain prob. distributions. It usually refers to the cases when the dist^n of a "waiting time" until a certain event, doesn't depend on how much time has elapsed already.

eg.:- exponential and geometric distributions.

Proof:- Let $G(y)$ be the CDF of y.

$$G_Y(y) = P(Y \le y).$$
$$= P(\phi(x) \le y).$$
$$= P(X \le \phi^{-1}(y)).$$

$\phi(x)$ is increasing f^n.

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

$$\boxed{g_Y(y) = F_X'(\phi^{-1}(y)) \cdot \frac{d}{dy}\phi^{-1}(y)}$$

If ϕ is decreasing f^n.

$$G_Y(y) = P(Y \le y).$$
$$= P(\phi(x) \le y)$$
$$= P(X \ge \phi^{-1}(y)).$$
$$= 1 - P(X \le \phi^{-1}(y))$$
$$= 1 - F_X(\phi^{-1}(y)).$$

$$\boxed{\therefore g_Y(y) = -F_X'(\phi^{-1}(y))\frac{d}{dy}\phi^{-1}(y)}.$$

Ex:- Let X be a continuous random variable with density f. Let's define $Y = a + bX$. Then what is the density of Y.

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

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

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

Scanned by CamScanner

$$f_X(x) = \begin{cases} \frac{2x}{R^2} &, 0 < x < R \\ 0 &, \text{o/w} \end{cases}.$$

$Y = \frac{X}{R}$ for $0 < Y < 1$. = Range of Y.

$g(y) = \frac{1}{|b|} \cdot f\left(\frac{y-a}{b}\right)$ Here, $a=0$, $b = 1/R$.

$$= R \cdot 2 \frac{yR}{R^2}$$

$$\begin{cases} g(y) = 2y &, 0 < y < 1. \\ 0 &, \text{o/w} \end{cases}.$$

Symmetric densities :-

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

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

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

Ex:- If X is a symmetric random variable with CDF $F_X(x)$. Then $F_X(0) = 1/2$.

$$\boxed{F_X(-x) = 1 - F_X(x) \quad \forall x \in \mathbb{R}.}$$

Next lecture is at the back.

Continue with Prof. Somesh's class :-

For, $\beta_2 = \left(\frac{\mu_4}{\mu_2^2} - 3\right) \rightarrow$ measure of peakedness.

$\beta_2 = 0 \rightarrow$ normal peak.

$\beta_2 > 0 \rightarrow$ leptokurtic / high peak.

$\beta_3 < 0 \rightarrow$ platykurtic / low peak.

Moment Generating Funcⁿ :-

For any random variable X, the moment generating f^n is defined as $M_X(t) = E(e^{tX})$. provided it exists, $t \in R$.

Consider,

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots$$

If expectation exists, we can take term by term expectation to get;

$$M_X(t) = E(e^{tX}) = 1 + tE(X) + \frac{t^2E(X^2)}{2!} + \cdots$$

$$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \cdots$$

that is, in the expansion on

Imp. formula :- if $aX+b$ is a new r.v. from X,
then, $\rightarrow E(aX+b) = aE(X) + b$.
$\rightarrow \sigma^2_{aX+b} = a^2\sigma_X^2$

We also consider $g(t) = M_X(t)$

then,
$$g(0) = 1.$$
$$g'(0) = \mu_1' \qquad g''(0) = \mu_2'$$

In general,
$$\left.\frac{d^r}{dt^r} M_X(t)\right|_{t=0} = \mu_r' \qquad r = 1, 2, \ldots$$

If MGF exists, then all moments exist.

Rel^{n}shp b/w Central and Non Central moments.

$$\mu_k = E(X-\mu_1')^k$$
$$= E\left[X^k - \binom{k}{1} X^{k-1}\mu_1' + \binom{k}{2} X^{k-2}\mu_1'^2 + \cdots + (-1)^{k+1}\mu_1'^k\right]$$
$$= E\left[\mu_k' - \binom{k}{1}\mu_{k-1}'\mu_1' + \binom{k}{2}\mu_{k-2}'\mu_1'^2 + \cdots + (-1)^{k+1}\mu_1'^k\right]$$

In particular,

$$\mu_2 = E(X-\mu_1')^2$$
$$= E\left[X^2 - 2X\mu_1' + \mu_1'^2\right]$$
$$= \mu_2' - \mu_1'^2$$
$$Var(X) = E(X^2) - \{E(X)\}^2$$

Scanned by CamScanner

PAGE No. DATE / /

Quantiles of a population :-

A number Q_p satisfying,

$$P(X \le Q_p) \ge p \quad \text{and} \quad P(X \ge Q_p) \ge 1-p$$

$0 < p < 1.$

is called p^{th} quantile of $dist^n$ of X. $\boxed{P(X \le Q_p) = p}$.

If F is an absolutely continuous cdf, then

$F(Q_p) =$

i.e. there exists a unique quantile.

For, $p = 1/2$, Q_p is called a median and is denoted by M.

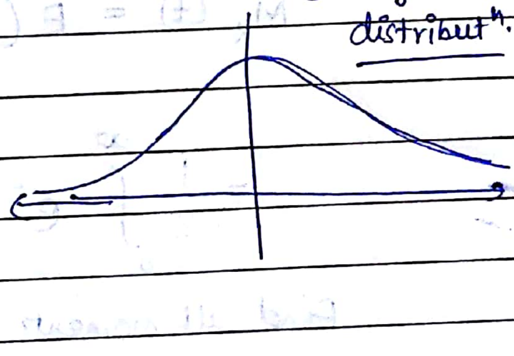
For, $p = 1/4, 1/2 \;\&\; 3/4$, $Q_1 = Q_2 = M$, Q_3 are called quartiles.

$Q_{1/100}, Q_{2/100}, \dots, Q_{99/100}$ are called percentiles.

Let, $f_X(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$, $-\infty < x < \infty$.

We want to find median.

Cauchy's $distribut^n$.



$F_X(x) = \frac{1}{\pi}(\tan^{-1}x + \pi/2)$.

So, $F_X(0) = 1/2$.

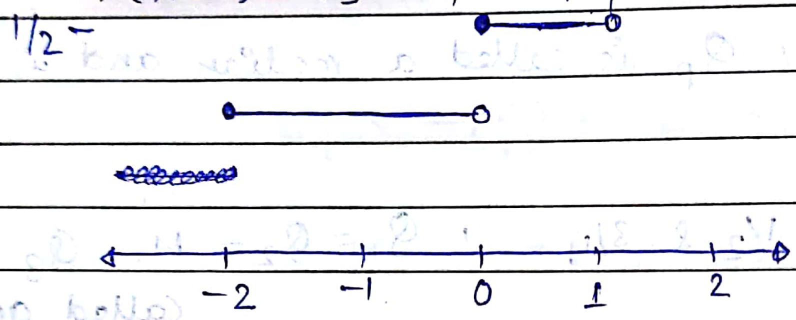
$\Rightarrow$ Median is 0.

$$F_X(Q_1) = 1/4 \Rightarrow \frac{1}{\pi}\left(\tan^{-1}x + \pi/2\right) = \frac{1}{4}$$

$$\Rightarrow \tan^{-1}x = -\pi/4 \;,\quad x = -1.$$

Example:- $P(X=-2) = 1/4$, $P(X=0) = 1/4$.

$P(X=1) = 1/3$, $P(X=2) = 1/6$.



$\therefore$ Any value b/w 0 and 1 can be the median.

$$f_X(x) = \begin{cases} \frac{1}{2} e^{-x/2} & , x > 0 \\ 0 & , x \le 0 \end{cases}.$$

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} \cdot \frac{1}{2} e^{-x/2}\, dx$$

$$= \frac{1}{2}\int_0^\infty e^{-x/2(1-2t)}\, dx = \frac{1}{1-2t} \;,\quad t < 1/2$$

Find all moments $\mu_1', \mu_2', \mu_3', \mu_4'$,
μ_2, μ_3, μ_4 from here,
β_1, β_2.

PAGE No.
DATE / /

Chebyshev's inequality :- Let X be a random variable with mean μ and variance σ^2.

Then for any $k > 0$:

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

Proof :- (Let us take continuous case).

Let X have pdf $f_X(x)$.

$$Var(X) = \sigma^2 = E(X-\mu)^2$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_X(x) \cdot dx.$$

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

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

Exercise :-

$$f_X(x) = \frac{1}{\beta}\left\{1 - \frac{|x-\alpha|}{\beta}\right\}.$$

$$\alpha-\beta < x < \alpha+\beta.$$

Find, μ_1', μ_2, μ_3, μ_4, Q_1, Q_2, Q_3 for this distⁿ.

Special Discrete distributions :-

① Degenerate Distⁿ :-

$$P(X=a) = 1.$$

for some point a.

$$E(X) = a, \quad V(X) = 0$$

② Discrete uniform distⁿ :-

X can take values $1, 2, \ldots, N$.

$$P(X=k) = \frac{1}{N} \qquad k = 1, 2, 3, \ldots, N.$$

$$E(X) = \sum_{k=1}^{N} \frac{k}{N} = \frac{N(N+1)}{2N} = \frac{N+1}{2}$$

$$E(X^2) = \sum_{j=1}^{N} \frac{j^2}{N} = \frac{(N+1)(2N+1)\cancel{N}}{6\cancel{N}}.$$

$$V(X) = E(X^2) - \{E(X)\}^2 = \frac{N^2-1}{12}$$

Moments of all order exists.

PAGE No.
DATE / /

$$M_X(t) = E(e^{tX}) = \sum_{j=1}^{N} e^{tj} \cdot \frac{1}{N}$$

$$= \begin{cases} \dfrac{e^t(e^{Nt}-1)}{N(e^t-1)}, & t \neq 0 \\ 1, & t = 0 \end{cases}$$

③ Bernoullian Trial :- If a trial results in two possible outcomes called 'success' & failure, it is called a Bernoullian trial.

$X(\text{success}) = 1, \quad X(\text{failure}) = 0.$

$P(X=1) = p, \quad P(X=0) = 1-p = q, \quad 0<p<1.$

This is called a Bernoulli distn.

$E(X) = p, \quad E(X^2) = p, \quad \mu_k' = p.$

$\mu_2 = Var(X) = p - p^2 = pq.$

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

$= (q + pe^t)$

④ Binomial distn :- Consider n independent trials of Bernoulli under identical condns with probability of success in each trial is p.

Let $X \longrightarrow$ number of successes in n trials.

Then, X can take values $0, 1, 2, \ldots, n$.

$$p_X(j) = P(X=j) = \binom{n}{j} p^j q^{n-j}, \quad j = 0, 1, 2, \ldots$$

$$\sum_{j=0}^{n} p_X(j) = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} = (q+p)^n = 1.$$

Scanned by CamScanner

$$\mu_1' = E(X)$$

$$= \sum_{j=0}^{n} j \binom{n}{j} p^j q^{n-j}$$

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

$$= \sum_{j=1}^{n} \frac{n!}{(j-1)!(n-j)!} p^j q^{n-j}$$

$$= np \sum_{j=1}^{n} \frac{(n-1)!}{(j-1)!(n-j)!} p^{j-1} \cdot q^{n-j}$$

$$= np(q+p)^{n-1} = np.$$

Now,

$$\mu_2' = E(X^2).$$

$$E\{X(X-1)\} + E(X) = E(X^2).$$

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

$$= \sum_{j=2}^{n} j(j-1) \frac{n!}{j!(n-j)!} p^j q^{n-j}$$

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

So, $E(X^2) = n(n-1)p^2 + np$

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

$$= np - np^2 = \underline{npq}.$$

in a binomial $dist^n$ $Var(X) < E(X)$.

$$S.d.(X) = \sqrt{npq}$$

Ex.:- Find $E(X^3)$, $E(X^4)$, μ_3, μ_4.

$$\mu_3 = np(1-p)(1-2p).$$

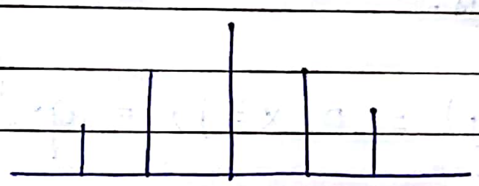
$$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{npq(1-2p)}{(npq)^{3/2}} = \frac{1-2p}{(npq)^{1/2}}$$

$= 0$ $p = 1/2 \rightarrow$ Symmetric.

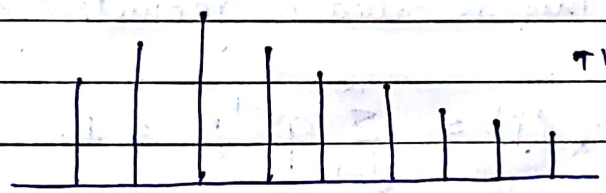
> 0 $p < 1/2 \rightarrow$ +ve skewed

< 0 $p > 1/2 \rightarrow$ -vely skewed.

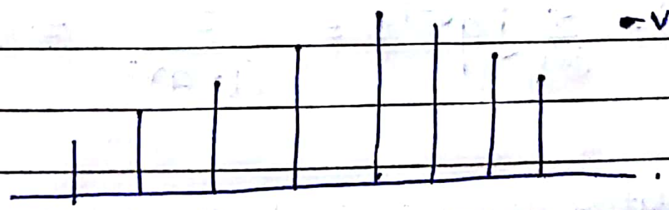
$p = 1/2$



$p < 1/2$ +vely skewed.



$p > 1/2$ -vely skewed.



$$\mu_4 = 3(npq)^2 + npq(1-6pq).$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{1-6pq}{npq}$$ of $pq = 1/6$.

>0 of $pq < 1/6$

<0 of $pq > 1/6$.

$$M_X(t) = \sum_{j=0}^{n} e^{tj} \binom{n}{j} p^j q^{n-j}$$

$$= \sum_{j=0}^{n} \binom{n}{j} (pe^t)^j q^{n-j}$$

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

MGF is unique for a distribution.

Suppose independent and identical Bernoullian trials are performed till we get the 1st success. Let X be the number of trials.

Then, $p_X(j) = P(X=j) = q^{j-1} p.$

This is called a geometric distribution.

$$\sum_{j=1}^{\infty} p_X(j) = \sum_{j=1}^{\infty} q^{j-1} p = 1.$$

$$\mu_1' = E(X) = \sum_{j=1}^{\infty} j q^{j-1} p = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

Calculating higher order factorial moments of this distribution we can use the following formula.

$$\frac{1}{(1-r)^{k+1}} = \sum_{j=k}^{\infty} \binom{j}{k} r^{j-k}$$

$$= \sum_{i=0}^{\infty} \binom{k+i}{k} r^i \qquad \underline{0 < r < 1}$$

Scanned by CamScanner

$\mu_2' = E(X^2) = \frac{q+1}{p^2}$

$Var(X) = \mu_2 = \mu_2' - \mu_1'^2 = \frac{q+1}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$

Moment generating f^n for a geometric $disb^n$:-

$M_X(t) = \sum_{j=1}^{\infty} e^{tj} q^{j-1} p = pe^t \sum_{j=1}^{\infty} (qe^t)^{j-1}$

$= \frac{pe^t}{1-qe^t} \qquad t < -\log_e q.$

$P(X > m) = \sum_{j=m+1}^{\infty} q^{j-1} p = q^m p + q^{m+1} p + \cdots\cdots$

$= q^m p \, (1 + q + q^2 + \cdots)$

$= \frac{q^m p}{1-q} = q^m$

P (success has not been observed till m trials).

$P(X > n+m \mid X > n) = \frac{P(\{X > n+m\} \cap \{X > n\})}{P(X>n)}.$

$= \frac{P(X > n+m)}{P(X > n)} = \frac{q^{n+m}}{q^n} = q^m = P(X > m).$

Memoryless property of geometric $distb^n$.

DATE / /

Ex:- Suppose independent tests are conducted on monkeys to develop a vaccine. If the prod.

Independent Bernoullian trials are performed under identical cond^ns till we get r^{th} success.

$$p_X(k) = P(X=k) = \binom{k-1}{r-1} q^{k-r} p^r$$

$$k = r, r+1, \ldots .$$

Negative Binomial distb^n.

Here, $E(X) = \frac{r}{p}$, $V(X) = \frac{rq}{p^2}$

$$M_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^r$$

$N \to$ Population size.

M ↓ type I people.

$(N-M)$ type II people.

A random sample of size n is selected from this pop^n. Let X be the number of items of type I in the sample.

Then,

$$p_X(x) = P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}.$$

Hypergeometric dist^n.

$x \le M$, $n-x \le N-M$

PAGE No.
DATE / /

$$\sum_{x=0}^{n} \binom{M}{x}\binom{N-M}{n-x} = \binom{N}{n}$$

$$\mu_1' = E(X) = \frac{Mn}{N} \quad , \quad E(X(X-1)) = \frac{M(M-1)\,n\,(n-1)}{N(N-1)}$$

$$E(X^2) = \frac{Mn\,(Mn - M - n - n + N)}{N(N-1)}$$

$$V(X) = \left(\frac{N-n}{N-1}\right)\frac{Mn}{N}\left(1-\frac{M}{N}\right)$$

Suppose we want to estimate the no. of tigers in a reserved forest.

Capture - Recapture technique :-

Theorem :- Let $X \sim$ Hypergeometric (M, N, n)
as, $M \to \infty$, $N \to \infty \ni \frac{M}{N} \to p$.

$$P(X=x) \to \binom{n}{x} p^x q^{n-x}$$

proof exercise.

Poisson process

The events occuring / being observed over time / area / space etc are referred to as being in a Poisson process of they satisfy the following three assumptions:

1. The no. of occurences during disjoint time intervals are independent.
   1D variable.

2. The probability of a single occurence during a small time interval is proportional to the length of the interval.
   $\boxed{P_1(h) = \lambda h} + o(h)$.

3. The probability of more than one occurence during a small time interval is negligible. $P(X(t) > 1)$
   $= P_2(t) + P_3(t) + \dots + \dots = o(h)$ → negligible

Let $X(t)$ = the no. of occurences in an interval of length t.

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

$\left(\frac{o(h)}{h} \to 0 \text{ as } h \to 0 \right)$

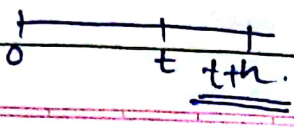
Under assumptions 1–3,

$$P_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!} \qquad n = 0, 1, 2, 3, \dots$$

Proof :- We will use principle of M.I.

To prove:- $P_0(t) = e^{-\lambda t}$, we consider

$P_0(t+h)$ = Probability of no occurence at interval $t+h$.



P (no occurrence in $(0, t+h]$)

$= P(\{\text{no occurrences in } (0,t]\} \cap \{\text{no occurrence in } (t, t+h]\})$

$= P(\{\text{no occurrence in } (0,t]\}) \cdot P(\text{no occurrence in } (t, t+h])$

$= P_0(t) \cdot P_0(h)$

From 3rd assumption,

$$P_0(t+h) = P_0(t) \cdot P_0(h) = P_0(t)\left(1 - \lambda h - o(h)\right)$$

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_0(t) - \frac{o(h)\, P_0(t)}{h}$$

Take lt as $h \to 0$.

$$P_0'(t) = -\lambda P_0(t)$$

$$P_0(t) = ce^{-\lambda t}$$

as, $P_0(0) = 1$. $\Rightarrow c = 1$

$$\boxed{\therefore P_0(t) = e^{-\lambda t}}$$

$P_1(t+h) = P$ (one occurrence in $(0, t+h]$).

$= P(\text{one occurrence in } (0, t]) \cap P(\text{no occurrence in } (t, t+h])$

$+ P(\text{one occurrence in } (t, t+h]) \cap P(\text{no occurrence in } (0, t])$.

$= P_1(t) \cdot P_0(h) + P_1(h) \cdot P_0(t)$.

$= P_1(t)\left[1 - \lambda h - o(h)\right] + e^{-\lambda t}(\lambda h + o(h))$

$$\Rightarrow \frac{P_1(t+h) - P_1(t)}{h} = -\lambda P_1(t) + \lambda e^{-\lambda t} - \frac{o(h)}{h} P_1(t) + \frac{o(h)}{h} e^{-\lambda t}$$

Taking lt as lt $h \to 0$,

$$P_1'(t) = -\lambda P_1(t) + \lambda e^{-\lambda t}$$

So the solⁿ is :-

$$P_1(t) = \lambda t e^{-\lambda t} + c \qquad P_1(0) = 0$$

$$P_1(t) = \lambda t e^{-\lambda t}$$

Let us assume $P_n(t)$ holds for all $n \leq k$.
Now for $n = k+1$.

$$P_{k+1}(t+h) = P_{k+1}(t) P_0(h) + P_k(t) P_1(h) + \sum_{j=2}^{k+1} P_{k+1-j}(t) \cdot P_j(h)$$

$$= P_{k+1}(t+h)(1-\lambda h - o(h)) + \frac{e^{-\lambda t}(\lambda t)^k}{k!}(\lambda(h)+o(h)) + \sum_{j=2}^{k+1} \frac{e^{-\lambda t}(\lambda t)^{k+1-j}}{(k+1-j)!} \underbrace{o(h)}_{\text{Negligible}}.$$

$$\frac{P_{k+1}(t+h) - P_{k+1}(t)}{h} = -\lambda P_{k+1}(t) + \frac{\lambda e^{-\lambda t}(\lambda t)^k}{k!} + \frac{o(h)}{h} \left(\text{terms containing } P_k(t) \right)$$

$$P_{k+1}'(t) = -\lambda P_{k+1}(t) + \frac{\lambda e^{-\lambda t}(\lambda t)^k}{k!}$$

The solution is :-

$$P_{k+1}(t) = \frac{e^{-\lambda t}(\lambda t)^{k+1}}{(k+1)!} + c \qquad P_{k+1}(0) = 0.$$

$c = 0$

Hence, proved.

Putting $\lambda t = \mu$, we call it a poisson distribution.

$$p_X(k) = P(X=k) = \frac{e^{-\mu}\mu^k}{k!} \qquad k = 0, 1, 2, \ldots$$

$$\sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} \frac{e^{-\mu}\mu^k}{k!} = e^{-\mu}.e^{\mu} = 1.$$

$$E(X) = \mu_1' = \sum_{k=0}^{\infty} k \frac{e^{-\mu}e^{\mu}}{k!} = \mu e^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!}$$

$$= \mu e^{-\mu}.e^{\mu} = \mu. = E(X)$$

$$E(X(X-1)) = \sum_{k=0}^{\infty}$$

$$= \mu^2$$

$$E(X^2) = E(X(X-1)) + E(X) = \mu^2 + \mu$$

$$Var(X) = E(X^2) - [E(X)]^2 = \mu^2 + \mu - \mu^2 = \mu.$$

So, in a Poisson distbⁿ mean and variance are the same.

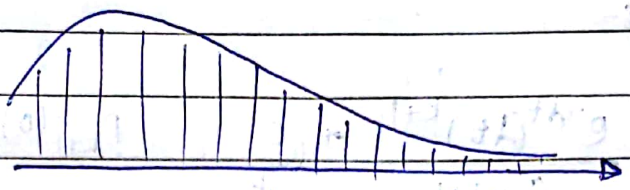
Find $\mu_3', \mu_4', \mu_3, \mu_4, \beta_1, \beta_2$.

$\mu_3' = \lambda^3 + 3\lambda^2 + \lambda$, $\mu_4' = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$.

$\mu_3 = \lambda$, $\mu_4 = \lambda + 3\lambda^2$

$$\beta_1 = \frac{1}{\sqrt{\lambda}} > 0 \quad \text{positively skewed.}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} - 3 = 1/\lambda > 0 \quad \text{leptokurtic.}$$



$$M_X(t) = E(e^{tX}) = \sum_{k=0}^{\infty} \frac{e^{tk} e^{-\mu} \mu^k}{k!}$$

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

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

Theorem :- Let $X \sim$ Binomial (n, p)

If $n \to \infty$ & $p \to 0$ such that $np \to \lambda$ then,

$$P(X=k) \to \frac{e^{-\lambda} \lambda^k}{k!} \qquad k = 0, 1, 2, 3, \ldots$$

There is one to one correspondence b/w MGF and distb's.

Proof:- Consider MGF of binomial distbn.

$$M_X(t) = (q + pe^t)^n \qquad np \approx \lambda.$$
$$p \approx \lambda/n.$$

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

$$= [1 + p(e^t - 1)]^n$$

$$\simeq \left[1 + \frac{\lambda}{n}(e^t - 1)\right]^n$$

$$\longrightarrow e^{\lambda(e^t - 1)}$$ which is mgf of $P(\lambda)$ distbn.

Give a direct proof.

PAGE No.

DATE / /

Suppose we are observing a Poisson process with rate λ. Let Y be the time of first occurrence. What is the distb^n of Y??

$$P(Y>t) = P(X(t)=0) \quad \text{same as.}$$

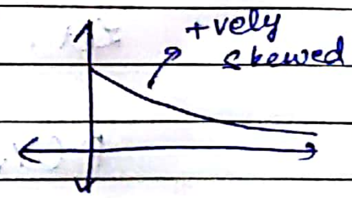
$$= \begin{cases} e^{-\lambda t}, & t>0 \\ 1, & t\le 0. \end{cases}$$

$$F_Y(t) = 1 - P(Y>t)$$

$$= \begin{cases} 0, & t\le 0 \\ 1-e^{-\lambda t}, & t>0 \end{cases}.$$

So, the pdf of Y is

$$f_Y(t) = \begin{cases} \lambda e^{-\lambda t}, & t\ge 0 \\ 0, & t\le 0. \end{cases}$$



$$E(Y^k) = \int_0^\infty t^k \lambda e^{-\lambda t}\, dt.$$

$$= \frac{\lambda \,\Gamma(k+1)}{\lambda^{k+1}} = \frac{k!}{\lambda^k} \quad \text{for } k=1,2,3,\ldots$$

$$\mu_1' = E(Y) = \frac{1}{\lambda}, \quad \mu_2' = \frac{2}{\lambda^2}$$

$$\mu_2 = V(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

$$\mu_3' = \frac{6}{\lambda^3}$$

$$\mu_3 = \frac{2}{\lambda^3}, \quad \mu_4 = \frac{9}{\lambda^4}.$$

$$\beta_1 = 2 > 0, \quad \beta_2 = 6 > 0$$ Independent of λ.

positively skewed & leptokurtic.

Scanned by CamScanner

DATE / /

MGF :- $E(e^{tY}) = \int_0^\infty e^{ty} \lambda e^{-\lambda y}\, dy.$

$$= \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

Memoryless property of exponential distb^n. most used in industry.

$$P(Y>t) = e^{-\lambda t}$$

$$P(Y>s+t \mid Y>s) = \frac{P(Y>s+t)}{P(Y>s)}$$

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

Suppose we are observing a Poisson process $X(t)$ with rate λ.

Let, Y_r denote the time of r^{th} occurrence.

$P(Y_r > t)$

$$= \sum_{j=0}^{r-1} P(X(t) \le j)$$

$$= \begin{cases} \displaystyle\sum_{j=0}^{r-1} \frac{e^{-\lambda t}\cdot(\lambda t)^j}{j!}, & t > 0 \\ 1, & t \le 0 \end{cases}$$

PAGE No.
DATE / /

$$F_{Y_r}(t) = 1 - P(Y_r > t)$$

$$= \begin{cases} 0 & , t \le 0 \\ 1 - \sum_{j=0}^{r-1} \frac{e^{-\lambda t}(\lambda t)^j}{j!} & , t > 0 \end{cases}$$

So, the pdf of Y_r is

$$f_{Y_r}(t) = -\frac{d}{dt}\left[e^{-\lambda t} + \lambda t e^{-\lambda t} + \frac{(\lambda t)^2}{2!} e^{-\lambda t} + \dots + \dots + \frac{(\lambda t)^{r-1} e^{-\lambda t}}{(r-1)!}\right]$$

$$f_{Y_r}(t) = \begin{cases} \frac{\lambda^r}{(r-1)!} e^{-\lambda t} t^{r-1} & , t > 0 \\ 0 & , t \le 0 \end{cases}$$

This is known as Erlang or Gamma distb^n.

General form of a Gamma (Erlang) distb^n is given as

$$f_X(x) = \frac{\lambda^r}{\Gamma r} e^{-\lambda x} x^{r-1} \quad , \; x > 0, \; \lambda > 0, \; r > 0.$$

$$\mu_k' = \int_0^\infty \frac{\lambda^r}{\Gamma r} e^{-\lambda x} x^{r+k-1} \, dx.$$

$$= \frac{\Gamma(k+r)}{\Gamma r} \cdot \frac{1}{\lambda^k} \quad , \; k = 1, 2, 3, \dots$$

$\mu_1' = E(X) = \dfrac{r}{\lambda}$ $\mu_2' = \dfrac{r(r+1)}{\lambda^2}$

$\mu_2 = Var(X) = \dfrac{r}{\lambda^2}$

MGF :-

$E(e^{(Y_r t)}) = \left(\dfrac{\lambda}{\lambda - t}\right)^r, \ t < \lambda.$