The theory of prob. deals with phenomenon that are random in nature which under some (statistical experiments) yields outcome indicating some pattern about the quantity of interest.

→ Random expt.

Random Expt.:

i) All possible outcomes are known in advance

ii) Any particular trial will yield an outcome that is known in advance

iii) Expt. can be repeated under identical conditions.

Sample Space (S or Ω):

Collection of all possible outcomes of RE.

Toss a coin Ten times

$$S = \{(e_1, e_2, e_3, \dots, e_{10}) \; ; \; e_i = H \text{ or } T, \; i = 1, 2, \dots, 10\}$$

Lifetime of a Machine

$$S = \{t \; ; \; t \geq 0\} \quad t \to \text{life in years}$$

$$S_1 = \{t \; ; \; 0 \leq t \leq \underset{\substack{\downarrow \\ \text{max life expectancy}}}{N}\}$$

Let A and B be events in S

$A \cup B \longrightarrow$ occurence of atleast one of events of A & B

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

$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \dots \cup A_n \longrightarrow$ occurence of atleast one of $A_1, A_2, \dots, A_n$

$A \cap B \rightarrow$ Simultaneous occurence of events A & B

$A^c \longrightarrow$ Event A doesn't occur.

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

Classical Def^n of Prob: (Laplace 1812)

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

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

* The theory of prob. has its origin in the games of chance which can be traced back atleast upto 17^{th} century.
* At the request of owners of some big gambling houses, Mathematicians like Laplace, Fermat, Pascal, De Moivre, developed the mathematical theory of prob.

④ Probability has found wide application in many field as tools of prob. and statistics can be used to explain and understand variation involved in areas like sciences & engg. application, clinical trials, life testing survival analysis, biology, computers, electronics etc.

7/1/2020

Relative frequency definition of Probability:

[Empirical def^n:]

Consider a RE with some sample space S.

Let expt. is conducted n no. of times and n_A be the no. of outcomes for event A.

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

the stabilization of this frequency for large sequence of trials is called statistical regularity

$$\text{So } P(A) = \frac{n_A}{n} \quad \text{as } n \longrightarrow \infty$$

Scanned with CamScanner

Axiomatic Definition of Probability

(A.H. Kolmogorov – Foundation of Prob. Theory)
1933

* most accepted definition of probability.

Let S be a sample space and P(A) denote the prob. of event A in S.

Then

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

Axiom II: $P(S) = 1$

$\rightarrow A \cap B = \phi$

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

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

Axiom III*: Let $A_1, A_2, \ldots, A_n$ be pair-wise disjoint events then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P(A_i) \qquad A_i \cap A_j = \phi \quad i,j = 1 \text{ to } n$$

Note:

(i) $P(\phi) = 0$

$$\phi \cup \phi^c = S$$

$$\underbrace{P(\phi) + P(\phi^c)}_{\text{By Axiom III}} = P(S) \Rightarrow P(\phi) + P(S) = P(S)$$

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

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

$$A \cup A^c = S$$

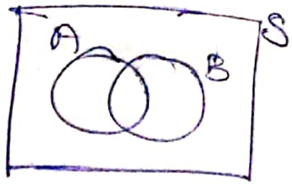
(iii)

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

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

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

~~Pf~~ $A = (A \cap B) \cup (B^c \cap A)$



Addition Formula for Prob.

Let A & B be 2 events in S. Then $P(A \cup B)$ is

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

$$\Rightarrow A \cup B = \underline{A \cap B^c} \cup \underline{A \cap B} \cup \underline{A^c \cap B}$$

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

$$= P(A) - \cancel{P(A \cap B)} + \cancel{P(A \cap B)} + P(B) - P(A \cap B)$$

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

General Formula for Add^n of Prob:

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

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

Pf: Math. Induction   Holds for $n=2$

~~consider $P(A \cup B \cup C)$~~

assume it to be true for $n=m$

$P\left(\bigcup_{i=1}^{m} A_i \cup A_{m+1}\right)$ show it true for $n=m+1$

Note:

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

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

Ex: Three popular option of a car model are

A: automatic B: new V_6-engine C: Air conditioning.

we have following data $P(A) = 0.7$ $P(B) = 0.75$

$P(C) = 0.8$ $P(A \cup B) = 0.8$ $P(A \cup C) = 0.85$ $P(B \cup C) = 0.9$

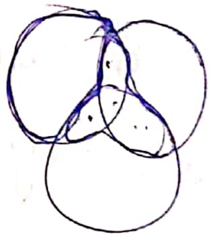
$P(A \cup B \cup C) = 0.95$

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

A: $P(A \cap B) = 0.65$

$(B \cap C) = 0.65$

$A \cap C = 0.65$



$$\begin{array}{r} 2.25 \\ -1.95 \\ \hline \end{array}$$

$0.95 = 0.3 + P(ABC)$

$P(ABC) = 0.65$

$0.95 - 0.65 - \cancel{0.65} - \cancel{0.65} + \cancel{1.3}$

$\boxed{0.3}$

8.1.2020

Eg: Toss a fair die twice

Define events

A: {atleast one six}

$P(A) = \frac{11}{36}$

event B : {sum of faces ≥ 10}

When this event is known then

what is the $P(A) = 5/6$

Conditional Probability:

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

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0 \qquad (1)$$

If $A \subset B \Rightarrow P(A|B) = \frac{P(A)}{P(B)} > P(A)$

$A \supset B \Rightarrow P(A|B) = 1$

(1) is a valid prob. func

$P(\cdot \mid B)$ is proper , $[0,1)$

some event given B

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

Axiom II: $P(S \mid B) = 1$

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

then $P(A \cup C \mid B) = P(A \mid B) + P(C \mid B)$

Multiplication Rule:

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

$$\parallel$$

$$P(A) P(B \mid A)$$

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

General Multiplication Formula:

Let $A_1, A_2, \dots\dots, A_n$ be n events in S. Then

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

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

Independent Events :

Two events A and B are independent

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

$\Rightarrow$ $P(A|B) = P(A)$ → here $P(B) \neq 0$ } hence we use

$P(B|A) = P(B)$ } $P(A) \neq 0$

⊛ Disjoint events are always dependent.

Q) Toss a coin 3 times:

A: Head on first count → $P(A) = \frac{1}{2}$

B: second → $P(B) = \frac{1}{2}$

C: third → $P(C) = \frac{1}{2}$

Are A,B,C independent?

⊛ In order to verify whether A,B,C are indep. we need to check

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

$P(A \cap C) = P(A) \cdot P(C)$ &

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

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

Q) Toss a coin twice :

A : head on 1^{st} try

B : 2^{nd}

C : same result

Are A, B, C independent event?

(*) Now consider 'n' events

$A_1, A_2, \ldots, A_n$. Then we have following conditions.

$\binom{n}{2}$ $P(A_i \cap A_j) = P(A_i)P(A_j)$ $\quad i<j = 1,2,\ldots,n$

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

$\vdots$

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

$\downarrow$

$2^n - n - 1$

9.1.2020

Total Prob. Result:

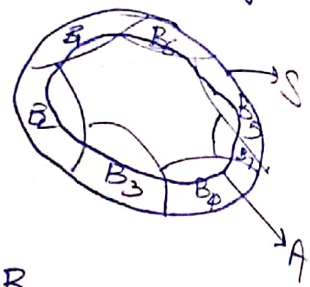
Let $B_1, B_2, \ldots, B_n$ be 'n' events of a sample space which form a partition of the given sample space. Now let A be any event in S, then

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

Pf: Since $B_1, B_2, \ldots B_n$ form a partition of S, hence these events are pairwise disjoint and they satisfy

$B_i \cap B_j = \phi$

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



we have

$$A = A\cap B_1 \cup A\cap B_2 \cup \ldots\ldots \cup A\cap B_n$$

Note that

$A\cap B_i$ and $A \cap B_j$ are pairwise disjoint

$i, j = 1, 2, \ldots, n$

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

$$= P(A\cap B_1) + P(A\cap B_2) + P(A\cap B_3) + \ldots + P(A\cap B_n)$$

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

Bayes Theorem (1764):

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

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

$i = 1, 2, \ldots, n$

Eg: Let 3 vendors A, B, C supply some product to IIT Patna. Suppose those 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 selected at random and tested.

[What is the prob. that it is defective?]

[Given that it is defective, what is the prob that it was supplied by i) A, ii) B, iii) C.

Eg: A binary communication channel carries data as one of 2 types of signals 0 and 1. Due to noise, a transmitted 0 is sometimes received as 1, and vice-versa. For a given channel, assume that a prob of 0.94 that a transmitted 0 is correctly received as 0 and a prob. 0.91 a transmitted 1 is correctly received as 1. Further assume that prob. of transmitting a 0 is 0.45. If signal is sent, Find the prob. that

i) a 1 is received

ii) a 1 was transmitted given that a 1 was received.

iii) an error occurs

A: i) $B^c = B^c \cap A \cup B^c \cap A^c$

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

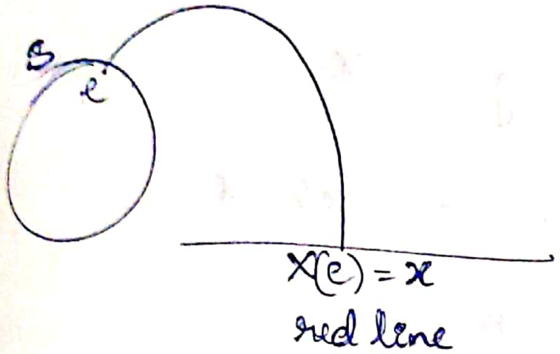
$$= P(A)\,P(B^c|A) + P(A^c)\,P(B^c|A^c)$$

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

13.1.2020

Random Variable

Consider a RE with sample space S. A random variable X is a ~~de~~ function defined over S which associates each outcome 'e' in S to a real number.

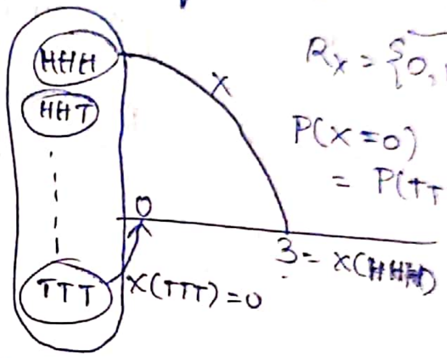


R_X : All possible values of X

After that we compute probability of each event in R_X

Ex: Toss a coin 3 times

X: no. of heads



$R_X = \{\overbrace{0, 1, 2, 3}^{x}\}$

$P(X=0)$
$= P(TTT) = \frac{1}{8}$

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

Cumulative Distribution Function: (CDF)

⇒ Let X be a RV with some prob. distⁿ then CDF of X is defined as

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

Properties of CDF:

i) $0 \leq F_X(x) \leq 1, \; \forall x \in R$

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

iii) $F_X(-\infty) = 0$

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

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

$F_X(x+h) = F_X(x)$ as $h \to 0^+$

vi) $P(X > x) = 1 - F_X(x)$

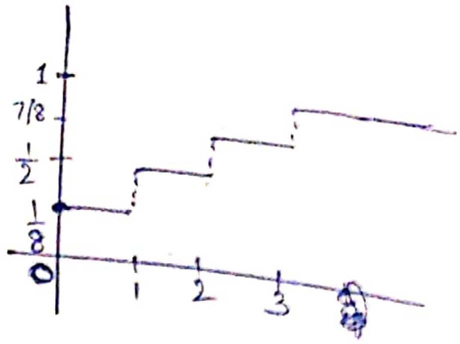
vii) $P(a < X \leq b) = F_X(b) - F_X(a)$

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

$$F_X(x) = P(X \leq x)$$

$$\begin{aligned} &= 0 && x < 0 \\ &= \frac{1}{8} && 0 \leq x < 1 \\ &= \frac{1}{8} + \frac{3}{8} = \frac{1}{2} && 1 \leq x < 2 \\ &= \frac{7}{8} && 2 \leq x < 3 \\ &= 1 && x \geq 3 \end{aligned}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{8} & 0 \le x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ \frac{7}{8} & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$



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

$$P(X=1) = F_X(1) - F_X(1^-)$$
$$= \frac{1}{2} - \frac{1}{8} = \frac{3}{8}$$

$$P(X=2) = F_X(2) - F_X(2^-)$$
$$= \frac{7}{8} - \frac{1}{2} = \frac{3}{8}$$

$$P(X=3) = F_X(3) - F_X(3^-)$$
$$= 1 - \frac{7}{8} = \frac{1}{8}$$

Ex:

Toss a fair die twice

X: Sum on upper faces

$R_X = \{2, 3, 4, \dots, 12\}$

Ex: Toss a coin until a head

X: no. of tosses to get that head.

$X > 4$

Types of RV's:

i) Discrete RV's

ii) Continuous RV's

iii) Mixed RV's

Discrete RV:

A RV X is said to be discrete if its range space contains atmost countably infinite no. of elements like $x_1, x_2, x_3, ---, x_i, ---$

Now let P_i be a number denoting the prob. of event $X=x_i$

i.e, $P_i = P(X=x_i) = \underbrace{P_X(x_i)}_{\text{short notation}}$

Then the collection $\{P_i \,;\, i \in \mathbb{N}\}$ is referred as the prob. mass funcⁿ (PMF) of X provided.

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

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

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

Q) 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

X	0	1	2
P_X	$\frac{{}^7C_2}{{}^{10}C_2}$	$\frac{{}^7C_1 \cdot {}^3C_1}{{}^{10}C_2}$	$\frac{{}^3C_2}{{}^{10}C_2}$
	$\frac{21}{45}$	$\frac{21}{45}$	$\frac{3}{45}$

	0	1	2
F_X	$\frac{21}{45}$	$\frac{42}{45}$	1

14.1.2020

Continuous RV:

A RV X is said to be Continuous if its CDF is defined as

$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt} \Rightarrow \frac{d}{dx} F_X(x) = f_X(x)$$

where $f_X(x)$ is known as probability density function (PDF) of X and it satisfies the following two properties $f_X(x)$

i) $\boxed{f_X(x) \geq 0 \quad \forall\, x \in R}$

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

$$P(a < x \leq b) = \int_a^b f_X(x)\,dx = P(a \leq x \leq b) = P(a < x < b)$$

$$= P(a \leq x < b)$$

$$\shortparallel$$

$$F_X(b) - F_X(a)$$

⊛ If X is continuous for any constant c, $\boxed{P(X=c) = 0}$

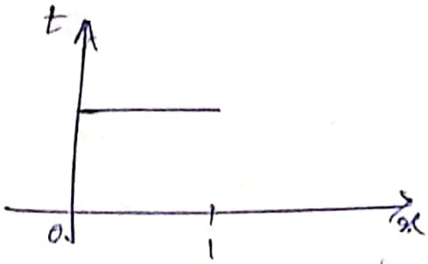
$\left.\begin{array}{l} a \leq x < b \\ a < x \leq b \\ a < x < b \\ a \leq x \leq b \end{array}\right\}$ if X is continuous all are equivalent.

* CDF of a continuous RV, X is continuous.

* $$\frac{dF_X(x)}{dx} = f_X(x)$$

Ex:

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$



$$P(0 < x < \tfrac{1}{2}) = \int_0^{1/2} dx = 1/2$$

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

Ex:

$$f_X(x) = \begin{cases} \dfrac{x}{2} & 0 \le x < 1 \\ \dfrac{k}{2} & 1 \le x < 2 \\ \dfrac{3-x}{2} & 2 \le x < 3 \\ 0 & \text{elsewhere} \end{cases}$$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 = \left.\frac{x^2}{4}\right|_0^1 + \left.\frac{kx}{2}\right|_1^2 + \left.\frac{3x}{2} - \frac{x^2}{4}\right|_2^3$$

$$\frac{1}{4} + \frac{k}{2} + \frac{3}{2} \cdot \frac{3}{2} - \frac{5}{4} = 1$$

$$\frac{k}{2} = \frac{1}{2} \Rightarrow \boxed{k = 1}$$

Find k so that $f_X(x)$ is a pdf. Also compute CDF of X.

$$\frac{15 - \frac{25}{4} - 5}{4} - \frac{1}{4}$$

To find k use the relation $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

A: $\int_0^1 \frac{x}{2}dx + \int_1^2 \frac{k}{2}dx + \int_2^3 \frac{3-x}{2}dx = 1$

$\frac{1}{4} + \frac{k}{2} + \frac{3}{2} - \frac{\cancel{5}1}{\cancel{4}} = 1$

$k = 1$

$$\Rightarrow f_X(x) = \begin{cases} \frac{x}{2} & 0 \le x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ \frac{3-x}{2} & 2 \le x < 3 \\ 0 & \text{otherwise} \end{cases}$$

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

$$F_X(x) = P(X \le x) = 0 \qquad x < 0$$

$$= \int_0^x \frac{t}{2}dt = \frac{x^2}{4} \qquad 0 \le x < 1$$

$$= \int_0^1 \frac{x}{2}dx + \int_1^x \frac{1}{2}dt = \frac{1}{4} + \frac{x-1}{2} \qquad 1 < x < 2$$

$$= \int_0^1 \frac{x}{2}dx + \int_1^2 \left(\frac{1}{2}\right)dx + \int_2^x \frac{3-t}{2}dt \qquad 2 < x < 3$$

$$= \frac{1}{4} + \frac{1}{2} + \frac{1}{2}\left(3t - \frac{t^2}{2}\right)_2^x = \frac{3x}{2} - \frac{x^2}{4} - \frac{5}{4}$$

$$1 \qquad x \ge 3$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2/4 & 0 < x < 1 \\ \frac{1}{4} + \frac{x-1}{2} & 1 < x < 2 \\ \frac{3x}{2} - \frac{x^2}{4} - \frac{5}{4} & 2 < x < 3 \\ 1 & x \geq 3 \end{cases}$$

$$P(\overbrace{X < 5/2}^{A} \mid \overbrace{X > 1}^{B}) = \frac{P(X < 5/2 \cap X > 1)}{P(X > 1)}$$

$$\underline{\underline{11/12}}$$

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

$$\int_1^\infty f_X(x)\,dx = \frac{1}{2}\int_1^2 dx + \int_2^3 \frac{3-x}{2}\,dx$$

<u>Mixed Type R.V:</u>

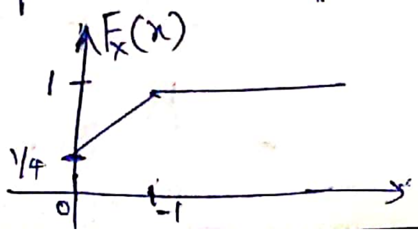
X: waiting time for service at counter

$P(X=0) = \frac{1}{4}$

$$f_X(x) = \begin{cases} \frac{3}{4} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & x = 0 \\ \frac{1}{4} + \frac{3}{4}x & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

$$P_X(0) + \int_0^1 \frac{3}{4}\,dx = \frac{1}{4} + \frac{3}{4} = 1$$



16·1·2020 Information provided by a RV:
Mean and Variance, Median, Quantile etc.

Mean: Mathematical Expectation (Average Value) of a RV X
Value of RV

Let X be a RV with some prob distⁿ, then expected value of X is defined as

$$\boxed{E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx}$$ [X is continuous with PDF $f_X(x)$]

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

If X is discrete with prob mass funcⁿ $P_X(x)$ then

$$E(X) = \sum_{x_i \in R_X} x_i P_X(x_i)$$

provided $\sum_{x_i} |x_i| P_X(x_i) < \infty$

X is continuous

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

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

⊛ Mean of a RV $\boxed{E(X) \text{ may not exist.}}$

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

||
discrete

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

Ex: Mean may not exist

Cauchy distⁿ $\quad f_X(x) = \frac{\beta}{\pi} \cdot \frac{1}{\beta^2 + (x-\alpha)^2}$

$\alpha, \beta \to$ parameters

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

Take $\alpha = 0 \cdot \beta = 1$

$$\underbrace{f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2} \quad -\infty < x < \infty}_{\text{Standard Cauchy}}$$

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

$E(|X|)$ does not exist

Ex: $f_X(x) = \begin{cases} \frac{1}{x^2} & x \geq 1 \\ 0 & \text{otherwise} \end{cases}$

Ex: $f_X(x) = 1 \quad 0 < x < 1$
$\quad = 0 \quad$ elsewhere

$$E(X) = \int_0^1 x f_X(x)\, dx$$
$$= \int_0^1 x\, dx = 1/2$$

Scanned with CamScanner

Ex: $f(x) = \frac{1}{2} e^{-x/2}, \quad x > 0$

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

Ex: X: no. of heads in three tosses

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

$$E(X) = \sum_{x>0}^{3} x \cdot p_X(x) = 0 \cdot P_X(0) + 1 \cdot P_X(1) + 2 P_X(2) + 3 \cdot P_X(3)$$

$$= 0 + \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2}$$

Variance:

Let X be a RV with some prob distⁿ

$$\boxed{V(X) = E(X - (EX))^2} = E[X^2 - 2XE(X) + (EX)^2]$$

$E(\cdot)$ is linear

$E(aX + b) = aE(X) + b$

Expected Value of const. is itself.

$$= E(X^2) - 2 \cdot E(X) \cdot E(X) + \cancel{EX} (EX)^2$$

$$\boxed{V(X) = E(X^2) - (EX)^2} \rightarrow \text{Simple formula}$$

always positive

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

$\downarrow$ continuous

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

Ex: X: no. of heads in 3 tosses

x	0	1	2	3
p	1/8	3/8	3/8	1/8

$$E(X^2) = \sum_{x=0}^{3} x^2 P_X(x)$$

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

$$E(X) = \frac{1}{2}$$

$$E(X^2) = \int_0^1 x^2\,dx = \frac{1}{3}$$

$$V(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} > 0$$

Ex: $P_X(0) = 3/4$

$f_X(x) = 1/4 \quad 0 < x < 1$

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

$$\boxed{E(X) = \frac{1}{8}}$$

20.1.2020

We know that,

$$\boxed{E(aX+b) = aE(x)+b}$$

(a, b given constants)

Now,

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

$\hookrightarrow E[aX+b - E(aX+b)]^2$

In particular,

$V(X+b) = V(X)$

$V(aX) = a^2 V(X)$

* Mean and Variance may or may not exists

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

Pf: $P(X \geq \alpha + x) \longrightarrow 1 - P(x < \alpha + x)$

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

If X is continuous

$F_X(\alpha - x) = 1 - F_X(\alpha + x)$

$\Rightarrow f_X(\alpha - x) = f_X(\alpha + x)$ $\forall x \in \mathbb{R}$

Spl. Case: If $\alpha = 0 \Rightarrow f_X(-x) = f_X(x)$ $\forall x \in \mathbb{R}$

Eg: Cauchy distⁿ: $f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2} \quad -\infty < x < \infty$

it is symmetric about $x = 0$. (given)

Consider Cauchy general form $\frac{\beta}{\pi}\frac{1}{\beta^2 + (x-\alpha)^2}$.

This distⁿ is symmetric about α.

Eg: $P_X(-1) = \frac{1}{4} = P_X(1)$, $P_X'(0) = \frac{1}{2}$

This distⁿ is symmetric about 0.

$$P(X \geq x) = P(X \leq -x) \quad \forall\, x \in \mathbb{R}$$

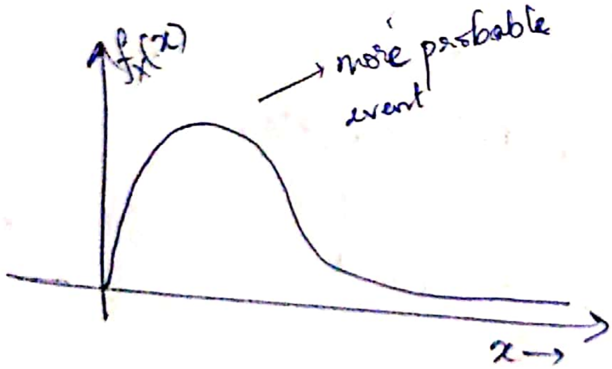
Coefficient of Skewness:

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

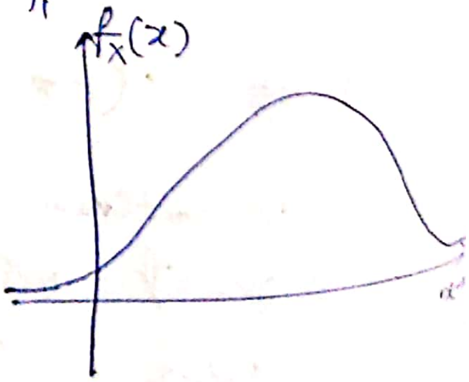
$$= \begin{cases} 0, & \text{symmetric} \\ >0, & \text{positively skewed} \\ <0, & \text{negatively skewed} \end{cases}$$

Standard Deviation: $s.d(X) = \sqrt{V(X)}$

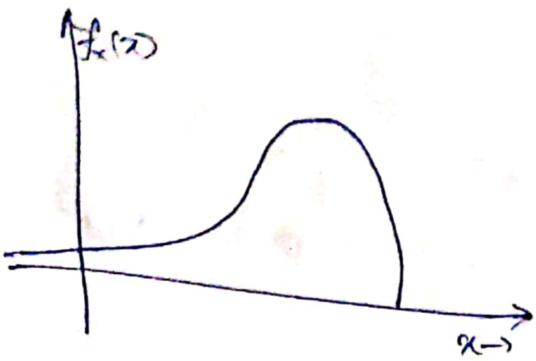
$\beta_1 > 0$



$\beta_1 = 0$



$\beta_1 < 0$



Quantiles: Let X be a RV with some prob. distⁿ

A number x_p is said to be p^{th} quantile of this distⁿ if,

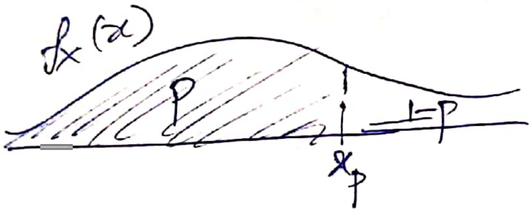
$\boxed{P(X \le x_p) \ge p}$ $\boxed{P(X \ge x_p) \ge 1-p}$

$0 < p < 1$

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

Let X be continuous, then

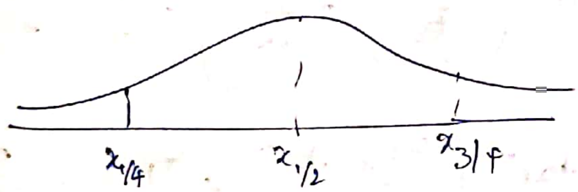
$\boxed{F_X(x_p) = p}$



p% will lie left of x_p

Consider $x_{1/4}, x_{1/2}, x_{3/4}$ quantiles of the dist. $f_X(x)$

$p = \frac{1}{2}$ or $x_{1/2}$ is the median of dist.



$x_{1/2}$ is another measure of central tendency & it always exists unlike mean.

④ Range of distⁿ $= x_{3/4} - x_{1/4}$

Eg: $f_X(x) = \frac{\beta}{\pi} \frac{1}{\beta^2+(x-\alpha)^2}$ $\quad -\infty < x < \infty$, $\alpha \in \mathbb{R}$, $\beta > 0$

Find quartiles of this distn.

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

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

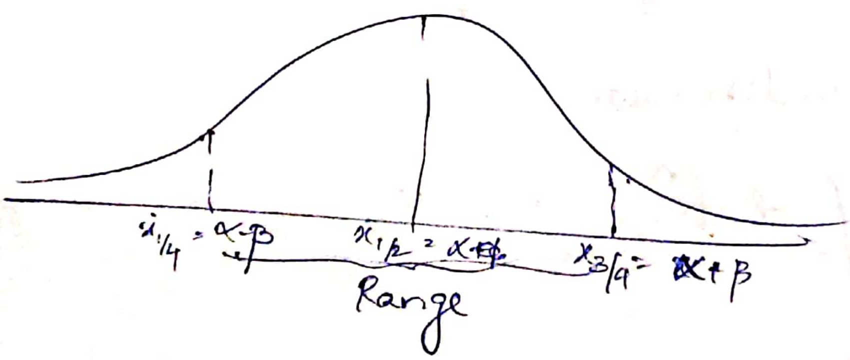
⊛ Range $= x_{3/4} - x_{1/4}$

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

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

$x_{3/4}$, $\frac{1}{\pi}\left[\tan^{-1}\left(\frac{x_{3/4}-\alpha}{\beta}\right) + \frac{\pi}{2}\right] = \frac{3}{4}$

$\Rightarrow \underline{x_{3/4} = \alpha+\beta}$



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

$P_X(1) = \frac{1}{6}$ $\quad P_X(2) = \frac{1}{3}$

$\Rightarrow$ Find median of this distⁿ

$\Rightarrow$ we need $x_{1/2}$ such that

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

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

Ans: Observe that $X=0$ satisfies the above 2 eqⁿs

$$P(x \geq 0) = \frac{1}{2} + \frac{1}{5} \geq \frac{1}{2}$$

$$P(x \leq 0) = \frac{1}{2}$$

Observe for $X=1$

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

$$P(X \leq 1) = \frac{1}{2} + \frac{1}{6} \geq \frac{1}{2}$$

For discrete case, more than 1 median can exists

(Here $X \in [0,1]$ are all medians, ∞^{ly} many median here)

(for discrete case)

For continuous case only one median possible.

21.1.2020

Moment Generating Function (MGF):

Suppose X is a RV with a given prob. distⁿ. Then MGF of X is given by.

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

discrete

$$\sum_{x \in R_X} e^{tx_i} P_X(x_i)$$

cont.

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

* $E(X)$ may not always exist

  ↳ $M_X(t)$ may not exist.

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

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

  $a, b \to$ constants

* $$\left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0} = E(X^n)$$

  $n = 1, 2, 3, \ldots$

Scanned with CamScanner

Discrete Case

eg:

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

Calculated before

$\underline{E(X) = 3/2}$

$\Downarrow$

$$M_X(t) = E[e^{Xt}]$$

$$= \sum_{x=0}^{3} e^{tx} P_X(x) = 1. P_X(0) + e^t. P_X(1) + e^{2t}. P_X(2) + e^{3t}. P_X(3)$$

$$\boxed{M_X(t) = \frac{1}{8} + \frac{3e^t}{8} + \frac{3e^{2t}}{8} + \frac{e^{3t}}{8}}$$

$$M_X'(t) = \frac{3}{8}e^t + \frac{3}{4}e^{2t} + \frac{3e^{3t}}{8}$$

$$M_X'(0) = \frac{3}{8} + \frac{3}{4} + \frac{3}{8} = \frac{3}{2} \longrightarrow E(X)$$

$\Uparrow$

Continuous Case

$$f_X(x) = \begin{cases} 3e^{-3x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

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

$$= 3\int_0^\infty e^{-(3-t)}\,dx$$

$$= \frac{3}{3-t}, \quad t < 3$$

$$E(X) = 3\int_0^\infty xe^{-3x}\,dx$$

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

$$V(X) = E(X^2) - (E(X))^2 = M_X''(0) - (M_X'(0))^2$$

⊛ For any RV X,

CDF, Mean, Variance, β_1, β_2. Quantiles (Median), Mode, MGF

⊛ MGF

Prob. Inequalities:

① Markov Inequality: Let X be a non-negative RV, with finite mean. If $a > 0$, a given constant then,

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

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

$$> \int_a^\infty x f_X(x)\,dx$$

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

② Tchebyshev Inequality:

Let X be a RV whose mean and ~~median~~ variance exist.

If $k>0$ is a given constant then,

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

μ → mean of X

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

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

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

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

$$\Rightarrow \sigma^2 \geq k^2 P(|X-\mu| \geq k)$$

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

$\int_{|x-\mu|<k} f_X(x) = P(|x-\mu|<k)$

$P(a<X<b) = \int_a^b f_X(x)\,dx$

Eg: $P_X(0) = 1-\frac{1}{k^2}$, $k>1$

$P_X(\pm 1) = \frac{1}{2k^2}$

Find an upper bound for the prob $P(|X| \geq 1)$ using Tchebyshev Ineq.

Ans: $E(X) = 0\left(1-\frac{1}{k^2}\right) + 1\cdot\frac{1}{2k^2} - 1\cdot\frac{1}{2k^2} = 0$

$E(X^2) = 0^2\left(1-\frac{1}{k^2}\right) + 1^2\cdot\frac{1}{2k^2} + 1^2\cdot\frac{1}{2k^2} = \frac{1}{k^2}$ $\quad V(X) = \frac{1}{k^2}$

$P(|X-E(X)| \geq 1) \leq \sigma^2$

$P(|X| \geq 1) \leq \frac{1}{k^2} \longrightarrow$ expected

$P(|X| \geq 1) = \frac{1}{k^2} \longrightarrow$ actual

Ex: $f_X(x) = \frac{1}{2\sqrt{3}} \qquad -\sqrt{3} < x < \sqrt{3}$

Find an upper bound for the prob. $P(|X| \geq \frac{3}{2})$

Compare with exact prob.

28.1.2020

Poisson Dist^n:

A RV X is said to ~~be~~ have a Poisson distribution if its PMF is of the following form,

$$P_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots \quad \lambda > 0$$

$$\boxed{X \sim P(X)}$$

- In Practice,
  Poisson dist^n is used to model rare events / phenomenon
- No. of defective products in a lot.
- No. of misprints in a book.
- No. of suicide reported.
- A particular type of car passing a road interval.

$$\sum_{x=0}^{\infty} P_X(x) = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda} \times e^{\lambda} = \underline{\underline{1}}$$

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

$$= e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!} \qquad x = y+1$$

$$= e^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^{y+1}}{y!} \Rightarrow \boxed{E(X) = \lambda}$$

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

$$= e^{-\lambda}\sum_{x=2}^{\infty} x(x-1)\frac{\lambda^x}{x!} + \lambda$$

$$= e^{-\lambda}\sum_{y=0}^{\infty} \cancel{y}\,\frac{\lambda^{y+2}}{y!} + \lambda$$

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

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

$$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2}$$

$$\Rightarrow \boxed{V(X) = \lambda}$$

Poisson dist^n is a unique dist^n with unique mean & variance.

MGF of a $P(\lambda)$ dist^n.

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

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

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

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

$\underline{t \neq 0}$

Scanned with CamScanner

$$\beta_1 = \frac{E(X-EX)^3}{\sigma^3} = \frac{E(X-\lambda)^3}{\lambda^{3/2}} = \frac{1}{\sqrt{\lambda}} > 0$$

Poisson approximation to the binomial distⁿ:

Let $X \sim Bin(n,p)$. Suppose n becomes very large and prob. of a success at each trial is small such that $np \to \lambda$, then

$$Bin(n,p) \longrightarrow P(\lambda) \quad \text{as } n \to \infty$$

$$\Rightarrow P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

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

$$= \frac{\lambda^x}{x!} \underbrace{\left[\frac{n(n-1)\cdots(n-x+1)}{n^x}\right] \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}}_{n\to\infty \text{ goes to } 1}$$

$$= \frac{\lambda^x}{x!} e^{-\lambda} \quad \text{as } n \to \infty$$

Ex: Suppose 1 in 5000 light bulbs are defective. Let X denote no. of defective light bulbs in a lot of 10,000 bulbs. What is the prob. that atleast 3 of them is defective

A. Approx. to Poisson,

$$X \sim P(\lambda) \quad \lambda = np = 2$$

$$\sim P(2)$$

$$P_X(x) = e^{-2} \frac{2^x}{x!}, \quad x = 0, 1, 2, \ldots$$

$$P(X \geq 3) = 1 - P_X(0) - P_X(1) - P_X(2)$$

$$= 1 - e^{-2} - 2e^{-2} - 2e^{-2}$$

30-11-2020

Geometric Distribution:

Conduct Bernoullian trials till first success is achieved

Define RV

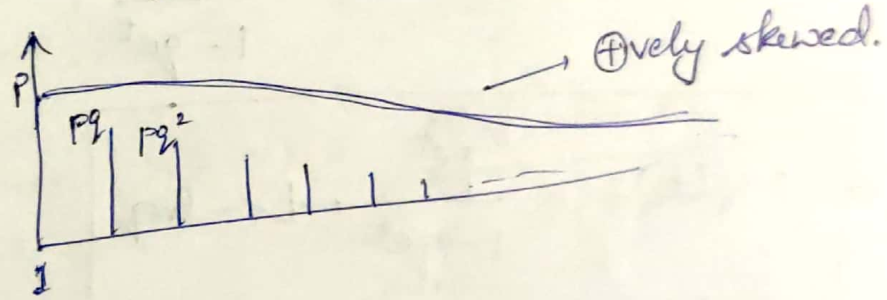
X: no. of trials required to get first success

range space, $R_X : 1, 2, 3, \dots$

$$P_X(x) = P(X=x) = q^{x-1}p \qquad x = 1,2,3,\dots \quad 0<p<1 \quad p+q=1$$

It is a proper prob. dist. because

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



$$E(X) = \sum_{x=1}^{\infty} x \cdot P_X(x) = \sum_{x=1}^{\infty} x q^{x-1} p$$

$$= p\,[1 + 2q + 3q^2 + \dots]$$

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

$$\boxed{X \sim Geo(p)}$$

Scanned with CamScanner

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

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

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

<u>MGF of a Geo(p) distⁿ:</u>

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

$$= p\left[e^{t} + e^{2t}q + e^{3t}q^2 + \dots \right]$$

$$= \frac{pe^t}{1-qe^t}, \quad qe^t < 1$$

$$\Rightarrow t < \ln\left(\frac{1}{q}\right)$$

$$\Rightarrow t < -\ln q$$

$$\boxed{\Rightarrow M_X(t) = \frac{pe^t}{1-qe^t}, \; t < -\ln q \\ = 1, \; t = 0}$$

$$g(t) = \frac{\frac{2}{3}e^t}{1-\frac{1}{3}e^t} \Rightarrow \boxed{P_X(x) = \left(\frac{1}{3}\right)^{x-1}\frac{2}{3}}$$

⊛ Let $X \sim Geo(P)$ $dist^n$, then find $P(X > b)$

Ai

$$\sum_{x=b+1}^{\infty} q^{x-1} p = \cancel{1 - \sum_{x=1}^{b} q^{x-1} p}$$

$$= \cancel{1 - q}$$

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

$$= \underline{\underline{q^b}}$$

Memoryless Property of a Geometric $Dist^n$:

Let $X \sim Geo(p)$ $dist^n$ and m & n given positive integers,

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

LHS $\Rightarrow$ $P(X > m+n \mid X > m) = \dfrac{P(X > m+n \cap X > m)}{P(X > m)}$

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

⊛ Geometric $dist^n$ is the unique discrete prob dist^n with memoryless property.

Alternative Representation:

Define

X: no. of trials preceding the first success

R_X : 0, 1, 2, ...

$P(X=x) = P_X(x) = q^x p \quad x = 0, 1, 2, \ldots$

$0 < p < 1$

$p + q = 1$

$E(X) = q/p$

$M_X(t) = \frac{p}{1-qe^t}, \; t < -\ln q$

$V(X) = q/p^2$

Q) Suppose independent trials are conducted on monkeys to develop a vaccine. If the prob. of success is 1/3 in each trial, then find the prob. that atleast 5 trials are reqd to get the first success.

A: X: no. of trials to get first success

$P_X(x) = \left(\frac{2}{3}\right)^{x-1} \frac{1}{3}, \quad x = 1, 2, \ldots$

$P(X \geq 5) = P(X > 4) = \left(\frac{2}{3}\right)^4$

Inverse Negative Binomial Dist^n:

Conduct Bernoullian Expt till a fixed 'r' no. of success are achieved

X: No. of trials read to get 'r' no. of successes

R_X: $r, r+1, r+2, r+3, \ldots$

$$P_X(x) = \binom{x-1}{r-1} p^{r-1} p q^{x-r}$$

$$= \binom{x-1}{r-1} p^r q^{x-r}$$

$x = r, r+1, r+2, \ldots$
$0 < p < 1$
$p + q = 1$

$$X \sim NB(r, p)$$

3/2/2020

$$r = 1 \Rightarrow NB(1, p) \approx Geo(p)$$

In fact ① in proper

$$p^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} q^{x-r}$$

put $x - r = t$

$$= p^r \sum_{t=0}^{\infty} \binom{t+r-1}{r-1} q^t = p^r \cdot \frac{1}{(1-q)^r} = \underline{\underline{1}}$$

$$\boxed{\sum_{t=0}^{\infty} \binom{t+r-1}{r-1} q^t = \frac{1}{(1-q)^r}}$$

$$E(X) = p^r \sum_{x=r}^{\infty} x \binom{x-1}{r-1} q^{x-r}$$

$$= p^r \sum_{x=r}^{\infty} \frac{x!}{(r-1)!\,(x-r)!} q^{x-r}$$

$$= rp^r \sum_{x=r}^{\infty} \binom{x}{r} q^{x-r} \qquad x-r=t$$

$$= rp^r \sum_{t=0}^{\infty} \binom{t+r}{t} q^t$$

$$= rp^r \frac{1}{(1-q)^{r+1}} = \frac{r}{p}$$

$$V(X) = \frac{rq}{p^2}$$

$$M_X(t) = E(e^{tX})$$

$$= p^r \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} q^{x-r} \qquad \text{put } x-r=y$$

$$= p^r \sum_{y=0}^{\infty} e^{t(y+r)} \binom{y+r-1}{r-1} q^y$$

$$= p^r e^{tr} \sum_{y=0}^{\infty} (qe^t)^y \binom{y+r-1}{r-1}$$

$$= \frac{p^r e^{tr}}{(1-qe^t)^r}, \qquad t < -\ln q$$

$$\Rightarrow \boxed{M_X(t) = \left[\frac{pe^t}{1-qe^t}\right]^r, \quad t < -\ln q}$$

Alternative form of an NB distⁿ:

X: no. of failures preceding a given 'r' no. of successes

$$P_X(x) = \binom{x+r-1}{x} p q^x p^{r-1}$$

$x = 0, 1, 2, \dots$

$0 < p < 1$

$p + q = 1$

$$\Rightarrow \boxed{P_X(x) = \binom{x+r-1}{x} q^x p^r}$$

$X \sim NB(r, p)$

$$p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} q^x = p^r \cdot \frac{1}{(1-q)^r} = 1$$

$$E(X) = \frac{rq}{p} \qquad V(X) = \frac{rq}{p^2}$$

$$M_X(t) = \frac{p^r}{(1-qe^t)^r}, \quad t < -\ln q$$

$$\boxed{\binom{x+r-1}{x} = (-1)^x \binom{-r}{x}}$$

Ex: A couple with 2 girls decided to keep having children until they have two boys. Prob of a male birth is 0.51. Find the prob. that couple will have atleast two more girls before completing a family. What is the expected size of such a family.

A: X: no. of girls before 2 boys

$X \sim NB(2, 0.5)$

$$P_X(x) = \binom{x+2-1}{2-1} (0.51)^2 (0.49)^x, \quad x = 0, 1, 2, \ldots$$

$$P(X \geq 2) = 1 - P(X < 2)$$

$$= 1 - P_X(0) - P_X(1)$$

$$\text{Expected Size} = 2 + 2 + E(X)$$

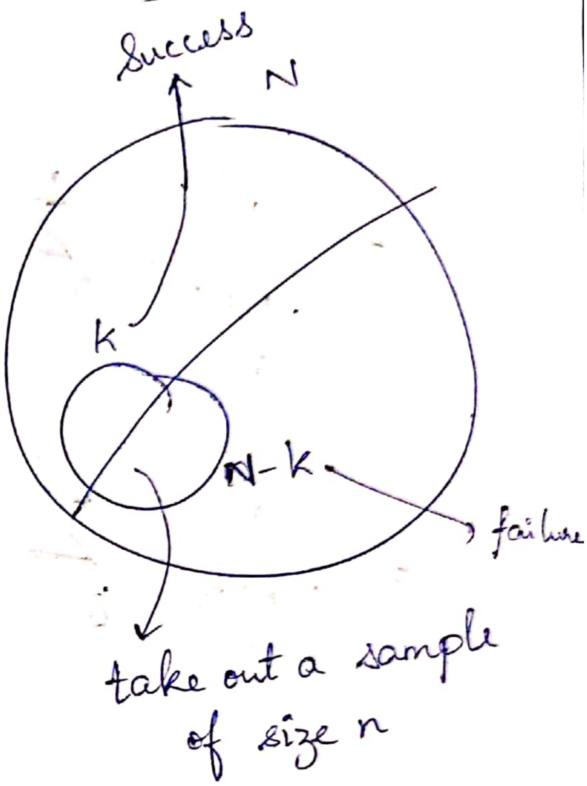
$$= 4 + \frac{2 \times 0.49}{0.51}$$

$$=$$

Hyper Geometric Distⁿ:

(Sampling without replacement)

X: no. of successes in a random sample of size n



$$\boxed{P_X(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \dots, n}$$

$$\boxed{X \sim (N, k, n)}$$

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

$$(1+x)^{N+M} = (1+x)^N (1+x)^M$$

$$\Rightarrow \sum_{i=0}^{N+M} \binom{N+M}{i} x^i = \left\{\sum_{j=0}^{N} \binom{N}{j} x^j\right\} \left\{\sum_{k=0}^{M} \binom{M}{k} x^k\right\}$$

$$= \left\{\binom{N}{0} + \binom{N}{1}x + \binom{N}{2}x^2 + \dots + \binom{N}{r}x^r + \dots + \binom{N}{N}x^N\right\}$$

$$\left\{\binom{M}{0} + \binom{M}{1}x + \binom{M}{2}x^2 + \dots + \binom{M}{r}x^r + \dots \binom{M}{M}x^M\right\}$$

Compare coeff. of x^r on both sides

$$\binom{N+M}{r} = \sum_{i=0}^{r} \binom{N}{i}\binom{M}{r-i}$$

$$E(X) = \sum_{x=0}^{n} \frac{x\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} \frac{k!}{(x-1)!(k-x)!}\binom{N-k}{n-x}$$

$$= \frac{1}{\binom{N}{n}} \sum_{y=0}^{n-1} \frac{k!}{y!(k-1-y)!}\binom{N-k}{n-1-y}$$

put $x-1=y$

$$= \frac{k}{\binom{N}{n}} \sum_{y=0}^{n-1} \frac{(k-1)!}{y!(k-1-y)!}\binom{N-k}{n-1-y}$$

$\to (N-1)-(k-1)$

$$= \frac{k}{\binom{N}{n}} \sum_{y=0}^{n-1} \binom{k-1}{y}\binom{N-k}{n-1-y}$$

$$= \frac{k}{\binom{N}{n}} \binom{N-1}{n-1} \Rightarrow E(X) = \frac{kn}{N}$$

$$V(X) = E(X(X-1)) + E(X) - (EX)^2$$

$$= \underline{\;?\;} + \frac{nk}{N} - \frac{n^2k^2}{N^2}$$

$$E\{X(X-1)\} = \frac{1}{\binom{N}{n}} \sum_{x=0}^{N} x(x-1)\binom{k}{x}\binom{N-k}{n-x}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=2}^{N} \frac{k!}{(x-2)!\,(k-x)!}\binom{N-k}{n-x}$$

$$= \frac{k(k-1)}{\binom{N}{n}} \sum_{y=0}^{N-2} \frac{(k-2)!}{(x-2)!\,(k-x)!}\binom{N-k}{n-x} \qquad \text{put } x-2=y$$

$$= \frac{k(k-1)}{\binom{N}{n}} \sum_{y=0}^{N-2} \binom{k-2}{y}\binom{N-k}{n-2-y}$$

$$= \frac{k(k-1)}{\binom{N}{n}} \binom{N-2}{n-2} = k(k-1)\frac{n(n-1)}{N(N-1)}$$

$$V(X) = E(X(X-1)) + E(X) - (EX)^2$$

$$= k(k-1)\frac{n(n-1)}{N(N-1)} + \frac{nk}{N} - \frac{n^2k^2}{N^2}$$

$$\boxed{V(X) = \frac{nk}{N}\cdot\frac{N-k}{N}\cdot\frac{N-n}{N-1}}$$

Note: MGF of a hypergeometric $dist^n$ doesn't exist in closed form expression.

Binomial Approximation to a HG $dist^n$:

Let $X \sim HG(N, k, n)$ $dist^n$.

If $N \to \infty$, $k \to \infty$ such that $\frac{k}{N} \to p$, then

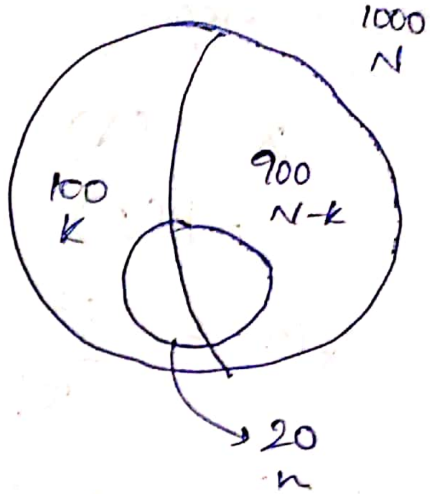
$$HG(N, k, n) \simeq Bin(n, p)$$

Equivalently,

$$\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \simeq \binom{n}{x} p^x (1-p)^{n-x}$$

Ex: Suppose 1000 bottles of a drink in filled by a machine during an hour. Each hour a sample of 20 bottles is selected. Let X be the no. of bottles that are underfilled. Also assume that each hour 100 under-filled bottles are produced by the machine. Find the prob. that atleast 3 underfilled bottles are there in the selected sample.

Ans: $X \sim HG(1000, 100, 20)$



$$P_X(x) = \frac{\binom{100}{x}\binom{900}{20-x}}{\binom{1000}{20}}$$

$x = 0, 1, 2, \ldots, n$

$$P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - P_X(0) - P_X(1) - P_X(2)$$

$$= 1 - \frac{1}{\binom{1000}{20}}\left\{\binom{100}{0}\binom{900}{20-0} + \binom{100}{1}\binom{900}{20-1} + \binom{100}{2}\binom{900}{20-2}\right\}$$

$$= \underline{\underline{0.3224}}$$

Approximation:

$$X \sim Bin(20, 0.1) \quad \left\{ P = \frac{K}{N} = \frac{100}{1000} = 0.1 \right\}$$

$$P_X(x) = \binom{20}{x}(0.1)^x (0.9)^{20-x}, \quad x = 0, 1, 2, \ldots 20$$

$$P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - \binom{20}{0}(0.9)^{20} - \binom{20}{1}(0.1)(0.9)^{19} - \binom{20}{1}(0.1)^2(0.9)^{18}$$

$$= \underline{\underline{0.3230}}$$

Another Approximation:

$$X \sim P(\lambda)$$
$$\sim P(2)$$

$$P_X(x) = \frac{e^{-2} 2^x}{x!}, \quad x = 0, 1, 2, \ldots$$

$$P(X \geq 3) = 1 - e^{-2} - 2e^{-2} - 2e^{-2}$$

$$= 1 - 5e^{-2} = 0.3233$$

Well known Continuous dist^n:

→ Continuous Uniform

→ Gamma, Beta

→ exponential

→ normal

→ lognormal

→ pareto

→ weibull

Continuous Uniform Distⁿ:

A RV X is said to have a continuous uniform dist if its PDF is given by

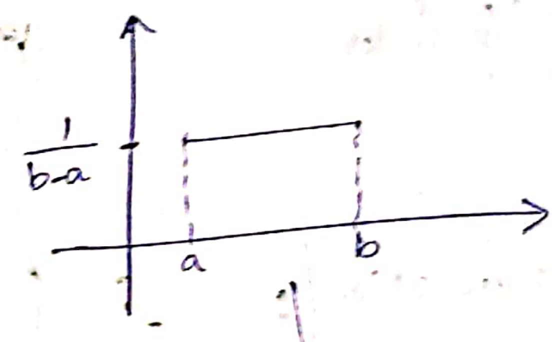
$$f_X(x,a,b) = \frac{1}{b-a}, \quad a \le x \le b$$
$$= 0 \quad ; \text{ otherwise}$$

$X \sim U[a,b]$

$\sim U[a,b)$

$X \sim U(-1,3)$

$X \sim U(0,1)$



$$F_X(x) = 0 \qquad x < a$$
$$= \frac{x-a}{b-a} \qquad a < x < b$$
$$= 1 \qquad x > b$$

$$F_X(x) = 0 \quad x < 0$$
$$= x \quad 0 < x < 1$$
$$= 1 \quad x > 1$$

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

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

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

$$E(X^2) = \frac{1}{b-a}\int_a^b x^2 dx = \frac{b^3 - a^3}{3(b-a)}$$

$$M_X(t) = \frac{1}{b-a}\int_a^b e^{tx} dx$$

$$= \frac{e^{tb} - e^{-ta}}{t(b-a)}, \quad t \ne 0$$

6.2.2020

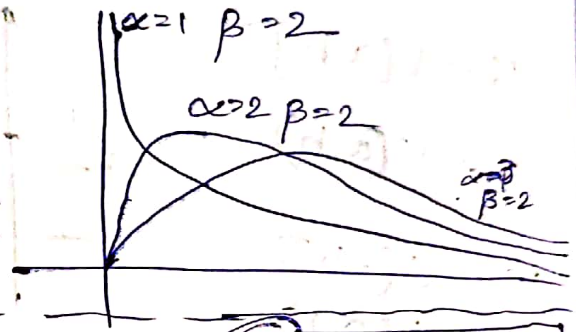
Gamma Dist^n:

A RV X is said to be a gamma dist^n provided its pdf is given by

$$\boxed{f_X(x,\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma\alpha} x^{\alpha-1} e^{-x/\beta}}, \quad \begin{array}{l} 0<x<\infty \\ \left.\begin{array}{l}0<\alpha<\infty \\ 0<\beta<\infty\end{array}\right\} \to \text{parameters}\end{array}$$

$$\boxed{X \sim G(\alpha,\beta)}$$

α → shape parameter; β → scale parameter



$\boxed{\Gamma\alpha = \text{Gamma function}}$:

* $\Gamma n = \int_0^\infty x^{n-1} e^{-x} dx, \quad n>0$

* $\boxed{\Gamma(n+1) = n\Gamma n}$

Eg: $\Gamma\frac{5}{2} = \frac{3}{2}\Gamma\frac{3}{2}$

$\Rightarrow \Gamma\frac{5}{2} = \frac{3}{2}\cdot\frac{1}{2}\Gamma\frac{1}{2} = \frac{3}{4}\sqrt{\pi}$

* $\boxed{\Gamma\frac{1}{2} = \sqrt{\pi}}$

$\boxed{\Gamma 1 = 1}$

* If n is a +ve integer,

$$\Gamma(n+1) = n!$$

Eg: $\Gamma 6 = 5!$

* Gamma dist^n is useful in modelling income, rainfall data, and is also widely used in reliability analysis.

To prove: $f_X(x, \alpha; \beta)$ is valid, now

$$\int_{-\infty}^{\infty} f_X(x, \alpha, \beta) = 1$$

Proof: $\int_{-\infty}^{\infty} f_X(x, \alpha, \beta)$ from def. $x \in (0, \infty)$

$$= \frac{1}{\beta^{\alpha} \Gamma\alpha} \int_{0}^{\infty} x^{\alpha-1} e^{-x/\beta} dx$$

put $\left(\frac{x}{\beta} = y\right)$

$$= \frac{1}{\beta^{\alpha} \Gamma\alpha} \int_{0}^{\infty} (y\beta)^{\alpha-1} e^{-y} \beta dy$$

$$= \frac{1}{\Gamma\alpha} \int_{0}^{\infty} y^{\alpha-1} e^{-y} dy$$

$$= \frac{1}{\Gamma\alpha} \Gamma\alpha = 1$$

$$\boxed{F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{\beta^{\alpha} \Gamma\alpha} \int_{0}^{x} t^{\alpha-1} e^{-t/\beta}, & x > 0 \end{cases}}$$

$$E(X^n) = \frac{1}{\beta^{\alpha} \Gamma\alpha} \int_{0}^{\infty} x^{n+\alpha-1} e^{-x/\beta} dx$$

$$= \frac{1}{\beta^{\alpha} \Gamma\alpha} \int_{0}^{\infty} (\beta y)^{n+\alpha-1} e^{-y} \beta dy$$

$$= \frac{\beta^{\alpha+n}}{\beta^{\alpha} \Gamma\alpha} \int_{0}^{\infty} y^{n+\alpha-1} e^{-y} dy = \boxed{\frac{\beta^n}{\Gamma\alpha} \Gamma(n+\alpha) = E(X^n)}$$

Scanned with CamScanner

$$E(X) = E(X^1) = \frac{\beta \Gamma\alpha+1}{\Gamma\alpha} = \boxed{\alpha\beta = E(X)}$$

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

$$V(X) = E(X^2) - (EX)^2 = \alpha(\alpha+1)\beta^2 - \alpha^2\beta^2$$

$$\boxed{V(X) = \alpha\beta^2}$$

<u>Mode of</u> X is solution of following eq^n.

$$\frac{d}{dx} f_X(x) = 0$$

$$\Rightarrow \quad x = \begin{cases} (\alpha-1)\beta, & \alpha > 1 \\ 0, & \alpha < 1 \end{cases}$$

<u>Median of</u> X is sol^n of $F_X(m) = \frac{1}{2}$

$$\frac{1}{\beta^\alpha \Gamma\alpha} \int_0^m t^{\alpha-1} e^{-t/\beta} dt = \frac{1}{2}$$

(and solve for m)

<u>MGF of Gamma dist^n</u>:

$$M_X(t) = E(e^{tX}) = \frac{1}{\beta^\alpha \Gamma\alpha} \int_0^\infty e^{tx} x^{\alpha-1} e^{-x/\beta} dx$$

$$= \frac{1}{\beta^\alpha \Gamma\alpha} \int_0^\infty e^{-x\left(\frac{1}{\beta}-t\right)} x^{\alpha-1} dx$$

put $x\left(\frac{1}{\beta}-t\right)$

$dx = \frac{dy}{\left(\frac{1}{\beta}-t\right)}$

$$\Rightarrow M_X(t) = \frac{1}{\beta^\alpha \Gamma\alpha} \int_0^\infty \left\{\frac{y}{\frac{1}{\beta}-t}\right\}^{\alpha-1} e^{-y} \frac{dy}{\left(\frac{1}{\beta}-t\right)}$$

$$= \frac{1}{\beta^\alpha \Gamma\alpha \left\{\frac{1}{\beta}-t\right\}^{\alpha}} \left(\int_0^\infty y^{\alpha-1} e^{-y} dy\right) \rightarrow \Gamma\alpha$$

$$= \frac{1}{(1-\beta t)^\alpha} \Rightarrow \boxed{M_X(t) = \frac{1}{(1-t\beta)^\alpha}}, \ t < \frac{1}{\beta}$$

$$\Rightarrow \boxed{M_X(t) = \begin{cases} \frac{1}{(1-t\beta)^\alpha} & t < \frac{1}{\beta} \\ 1 & t = 0 \end{cases}}$$

Eg: $M_X(t) = \frac{1}{(1-2t)^3} \quad t < \frac{1}{2}$

This is MGF of $G(3,2)$

Let $X \sim G(\alpha, \beta)$

Take a particular case

$\alpha = \frac{n}{2}$, $\beta = 2$

n is a positive integer,

then resulting prob. dist. is known as

$\chi^2(n)$ distn

$X \sim \chi^2(n) \quad f_X(x) = \frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma n/2} \quad \begin{array}{l} x > 0 \\ n \text{ is a } \\ \text{(+) integer} \end{array}$

Scanned with CamScanner

Beta distⁿ:

A RV X is said to have the distⁿ if its pdf is

$$f_X(x,\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \quad 0<x<1,\ \alpha>0,\ \beta>0$$

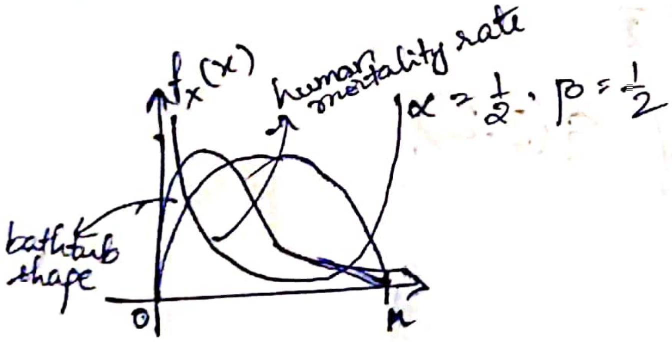
$B(\alpha,\beta)$ → beta func.

$\boxed{X \sim Beta(\alpha,\beta)}$

Beta funcⁿ:

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx \quad \alpha>0,\ \beta>0$$

$$\boxed{B(\alpha,\beta) = \frac{\Gamma\alpha\,\Gamma\beta}{\Gamma(\alpha+\beta)}}$$



Note: MGF of Beta func won't exists in closed form

$X \sim Beta(1,1) \approx U(0,1)$

↓

$$f_X(x,1,1) = \frac{1}{\Gamma 1}$$

$$\boxed{f_X(x,1,1) = 1} \quad 0<x<1$$

CDF is not in closed form

$$F_X(x) = \begin{cases} 0 & x<0 \\ \frac{1}{B(\alpha,\beta)} \int_0^x t^{\alpha-1}(1-t)^{\beta-1}\,dt & 0<x<1 \\ 1 & x \geq 1 \end{cases}$$

MGF not in closed form

$$E(X^n) = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{n+\alpha-1}(1-x)^{\beta-1}\,dx$$

$$\boxed{E(X^n) = \frac{B(n+\alpha,\beta)}{B(\alpha,\beta)}}$$

$$E(X) = \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)}$$

$$= \frac{\Gamma(\alpha+1)\cancel{\Gamma\beta}}{\Gamma\alpha+\beta+1} \times \frac{\Gamma\alpha+\beta}{\Gamma\alpha\,\cancel{\Gamma\beta}}$$

$$\boxed{E(X) = \frac{\alpha}{\alpha+\beta}}$$

$$E(X^2) = \frac{\Gamma\alpha+2\,\cancel{\Gamma\beta}}{\Gamma\alpha+\beta+2} \times \frac{\Gamma\alpha+\beta}{\Gamma\alpha\,\cancel{\Gamma\beta}}$$

$$E(X^2) = \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}$$

$$\Rightarrow E(X^2)-(EX)^2 = \frac{\alpha}{\alpha+\beta}\left[\frac{\alpha+1}{\alpha+\beta+1} - \frac{\alpha}{\alpha+\beta}\right]$$

Mode: $\dfrac{\alpha-1}{\alpha+\beta-2}$, $\alpha>1$, $\beta>1$.

$X \sim G(\alpha,\beta)$

Take $\alpha=\frac{n}{2}, \beta=2$

$X \Rightarrow \chi^2(n)$

One-Parameter Exponential Dist^n:

A RV X is said to have a one-parameter exponential dist^n if its pdf is

$$f_X(x,\beta)=\frac{1}{\beta}e^{-x/\beta}, \quad 0<x<\infty, \quad 0<\beta<\infty$$

$\boxed{X \sim Exp(\beta)}$

Useful in lifetime analysis

$$\frac{1}{\beta}\int_0^\infty e^{-x/\beta}dx = 1$$

$E(X)=\beta$

$V(X)=\beta^2$

$M_X(t)=\dfrac{1}{1-t\beta}, \quad t<\dfrac{1}{\beta}$

$F_X(x)=\frac{1}{\beta}\int_0^x e^{-x/\beta}dx$

$\boxed{F_X(x)=1-e^{-x/\beta}}$

$\Downarrow$

$P(X>s)=e^{-s/\beta}$

For median,

$F_X(m)=\frac{1}{2} \Rightarrow \boxed{m=\beta\ln(2)}$

Memoryless Property:

Let X be a RV s.t. $\boxed{X \sim exp(\beta)}$ distⁿ and $s>0, t>0$ be given constants, then $P(X > s+t \mid X > s) = P(X>t)$

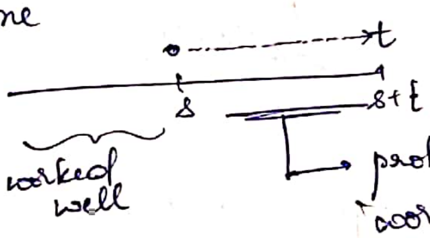
$$\frac{P(X > s+t \cap P(X > s))}{P(X > s)} = \frac{P(X > s+t)}{P(X>s)}$$

$$= \frac{e^{-\frac{s+t}{\beta}}}{e^{-s/\beta}} = e^{-t/\beta}$$

$$= P(X>t)$$

$$\Rightarrow \boxed{P(X > s+t \mid X > s) = P(X>t)}$$

X: lifetime



probability that it would work well for further time t.

$$\frac{\xi^{\alpha} x^{\alpha-1} e^{-\xi x}}{\Gamma\alpha}, \quad x>0, \ \xi>0, \ \alpha>0$$

Gamma PDF:

$$\frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma\alpha\, \beta^{\alpha}}$$

$$\alpha = 1 \Rightarrow \frac{1}{\beta} e^{-x/\beta}$$

$$\xi = \frac{1}{\beta}$$

$$\frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma\alpha}$$

$$E(X) = \alpha/\beta$$

$$V(X) = \alpha/\beta^2$$

Two-parameter exponential distn:

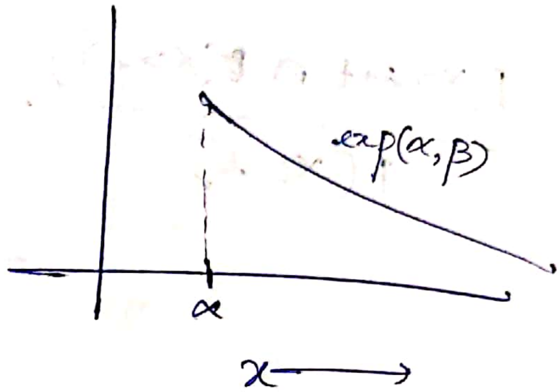
$$f_X(x,\alpha,\beta) = \frac{1}{\beta} e^{-\frac{(x-\alpha)}{\beta}}$$

$\alpha \le x < \alpha$

$\alpha \in \mathbb{R}$

$\beta > 0$

$X \sim exp(\alpha, \beta)$



$E(X) = \alpha + \beta$

$V(X) = \alpha^2$

$$F_X(x) = 1 - e^{-\frac{x-\alpha}{\beta}}$$

13.2.2020

Normal Distribution:

$$f_X(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-1}{2\pi^2}(x-\mu)^2}$$

$-\infty < x < \alpha$

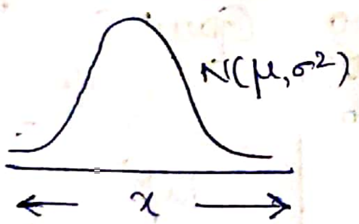
$-\infty < \mu < \infty$

$0 < \sigma < \infty$

$$\boxed{X \sim N(\mu, \sigma^2)}$$

$\mu \to$ location parameter

$\sigma \to$ scale parameter



No shape parameter

(shape of a $N(\mu,\sigma^2)$ is a bell-shape)

Probably the most widely used prob. distn in literators

Proof for Normal Dist^n to be proper:

$$\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\,dx$$

$\frac{x-\mu}{\sigma} = t$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\,dt$$

$\frac{t}{\sqrt{2}} = u$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2}\,du$$

$u^2 = k$

$2u\,du = dk$

$$\frac{2}{\sqrt{\pi}}\int_{0}^{\infty}\frac{e^{-k}}{2\sqrt{k}}\,dk = \frac{1}{\sqrt{\pi}}\int_0^\infty k^{\frac{1}{2}-1}e^{-k}\,dk = \frac{1}{\sqrt{\pi}}\times\sqrt{\pi}$$

$$\int_{-\infty}^{\infty} e^{-u^2}\,du \int_{-\infty}^{\infty} e^{-v^2}\,dv = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(u^2+v^2)}\,du\,dv$$

$$= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,r\,dr\,d\theta =$$

$2r\,dr = dk$

$$= \frac{1}{2}\int_0^{2\pi}\int_0^{\infty} e^{-t}\,dt\,d\theta = \frac{1}{2}\int_0^{2\pi} 1\,d\theta = \pi$$

$e^{-t}\Big|_{\infty}^{0}$

$$\Rightarrow \boxed{\int_{-\infty}^{\infty} e^{-u^2}\,du = \sqrt{\pi}}$$

$$\Rightarrow \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2}\,du = 1$$

Normal Dist^n is proper

Scanned with CamScanner

$$E(X) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} x\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad \frac{x-\mu}{\sigma} = t$$

$$= \frac{1}{\sqrt{2\pi}\,\cancel{\sigma}}\int_{-\infty}^{\infty} (\sigma t + \mu)\, e^{-\frac{t^2}{2}} dt$$

$$= \frac{1}{\sqrt{2}\sqrt{\pi}}\left[\underbrace{\int_{-\infty}^{\infty} \sigma t e^{-\frac{t^2}{2}} dt}_{\text{odd}} + \underbrace{\int_{-\infty}^{\infty} \mu e^{-\frac{t^2}{2}} dt}_{\text{even}}\right]$$

$$= \frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \frac{\mu}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2} du$$

$$\frac{t}{\sqrt{2}} = u$$

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

$$E(X^2) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad \frac{x-\mu}{\sigma} = t$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} (\mu + \sigma t)^2 e^{-\frac{t^2}{2}} dt$$

$$= \frac{1}{\sqrt{2\pi}}\left[\int_{-\infty}^{\infty}\left(\mu^2 e^{-\frac{t^2}{2}} + \overbrace{2\mu\sigma t e^{-\frac{t^2}{2}}}^{\text{odd}} + \sigma^2 t^2 e^{-\frac{t^2}{2}}\right) dt\right]$$

$$= \mu^2 + \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt \qquad u^2 = t,\quad du = \frac{dt}{2\sqrt{t}}$$

$$= \mu^2 + \frac{\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty} 2u^2 e^{-u^2} du$$

$$= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} \cancel{2} t e^{-t} \frac{dt}{\cancel{2}\sqrt{t}}$$

$$= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \int_0^\infty t^{\frac{3}{2}-1} e^{-t} dt$$

$$= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \times \Gamma\frac{3}{2} = \mu^2 + \frac{\cancel{2}\sigma^2}{\cancel{\sqrt{\pi}}} \times \frac{1}{\cancel{2}} \cancel{\sqrt{\pi}}$$

$$= \mu^2 + \sigma^2$$

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

$$\Rightarrow \boxed{V(X) = \sigma^2}$$

~~Note~~

Co-efficient of Skewness (β_1)

$\because$ Normal Distⁿ is symmetric about μ,

$$\Rightarrow \boxed{\beta_1 = 0}$$

MGF: $M_X(t) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{xt} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

$\frac{x-\mu}{\sigma} = k$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{(\mu+\sigma k)t} e^{-\frac{k^2}{2}} dk$$

$$\boxed{M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}}$$

Scanned with CamScanner

$$F_X(x) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\boxed{\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}} du}$$

$$= \int_{-\infty}^{z} \phi(u) du$$

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

$Z \sim N(0,1)$

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$$

$u \in \mathbb{R}$

Log Normal Distribution:

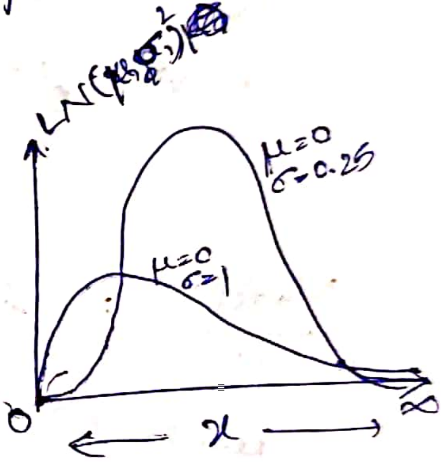
A RV X is referred as lognormal provided $\ln(X)$ is normally distributed.

The pdf of X is

$$f_X(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma x} e^{-\frac{1}{2\sigma^2}(\ln x-\mu)^2}, \quad \begin{array}{l} x>0 \\ \mu\in\mathbb{R} \\ \sigma>0 \end{array}$$

we showed that it is proper pdf.

$$F_X(x) = \Phi\left(\frac{\ln x-\mu}{\sigma}\right)$$



MGF doesn't exist

$$E(X) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_0^\infty x\cdot\frac{1}{x}\, e^{-\frac{1}{2\sigma^2}(\ln x-\mu)^2}\,dx$$

$$\frac{\ln x-\mu}{\sigma} = t$$

$$\frac{dx}{\sigma x} = dt$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2/2}\, e^{\sigma t}\,dt$$

$$= e^{\mu+\frac{\sigma^2}{2}}\underbrace{\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t-\sigma)^2}\,dt\right]}$$

$\to \int_{-\infty}^{\infty} f_X(x)\,dx$ of normal dist^n with $\mu=\sigma$, $\sigma=1$

$$\boxed{E(X) = e^{\mu+\frac{\sigma^2}{2}}}$$

Alternative Approach:

$$X \sim LN(\mu, \sigma^2)$$

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

$$\Rightarrow \quad X = e^Y$$

$$E[X^k] = E(e^{Yk}) = e^{k\mu + \frac{1}{2}\sigma^2 k^2}$$

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

$$\boxed{V(X) = e^{2\mu + 2\sigma^2} - e^{\mu + \frac{1}{2}\sigma^2}}$$

Mode(X) : Solve $\frac{d}{dx} f_X(x) = 0$ for x

$\parallel$

$e^{\mu - \sigma^2}$

Median(X) : $\cancel{F_X}$ $\Phi\left(\frac{\ln x - \mu}{\sigma}\right) = \frac{1}{2}$

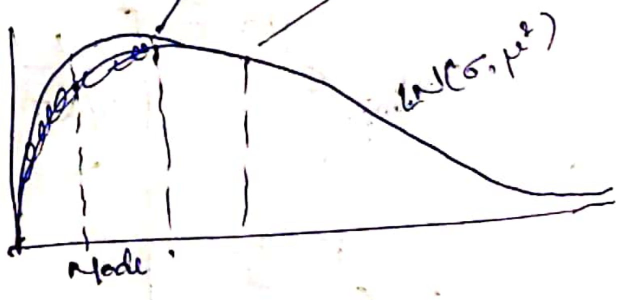
$$\Rightarrow \frac{\ln x - \mu}{\sigma} = 0 \Rightarrow \boxed{x = \cancel{\mu} e^{\mu}}$$

Mean: $e^{\mu + \frac{\sigma^2}{2}}$

Mode: $e^{\mu - \sigma^2}$

Median: e^{μ}

Mode < Median < ~~Mode~~ Mean



Q) Let $X \sim N(\mu, \mu^2)$, $\mu > 0$

Write the probability $P(X < -\mu \mid X < \mu)$ in terms of CDF of a standard normal variable

A: $P(X < -\mu \mid X < \mu) = \dfrac{P(X < -\mu)}{P(X < \mu)}$

$$= \frac{P\left(\frac{X-\mu}{\mu} < \frac{-\mu-\mu}{\mu}\right)}{P\left(\frac{X-\mu}{\mu} < \frac{\mu-\mu}{\mu}\right)}$$

$$= \frac{P(Z < -2)}{P(Z < 0)} = \frac{\Phi(-2)}{\Phi(0)}$$

$$= 2\Phi(-2)$$

Weibull distⁿ: (Swedish Physicist W. Weibull (1939)):

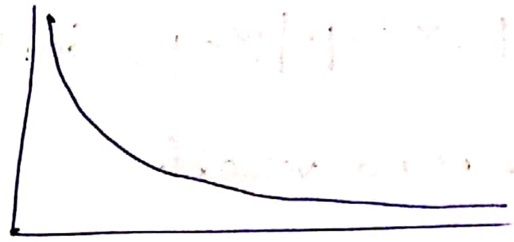
$$\boxed{f_X(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha},} \quad \begin{array}{l} x > 0 \\ \alpha > 0 \\ \beta > 0 \end{array}$$

$X \sim \text{Weibull}(\alpha, \beta)$ $\qquad F_X(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha}$

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

$$V(X) = \beta^2 \left\{ \Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right)^2 \right\}$$

Pareto distⁿ:



$$f_X(x) = \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}}, \quad x \geq \beta,\ \alpha > 0,\ \beta > 0$$

$$F_X(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}$$

$$E(X^n) = \beta^n \frac{\alpha}{\alpha - n}, \quad n < \alpha$$

20·2·2020

Transformation of a Random Variable

You are given a prob. dist. ~~X~~ of RV X.

Our interest is in studying probabilistic behaviour of some func. of X e.g.,

$$Y = g(X)$$

$$= \underbrace{|X|,\ X^2,\ aX+b,\ \ln X}_{\text{it can be anything (any function of } X)},\ e^X,\ \max(x,0) \ldots$$

Once you have your function of interest, you want to compute corresponding dists.

Some important methods to compute prob. dist^n of Y are:

i) CDF approach

ii) MGF approach

iii) Fundamental Theorem (for continuous case only)

CDF Approach:

Let $Y = aX + b$, $b \in R$
$a \neq 0$, any constant

$$F_Y(y) = P(Y \leq y)$$

$$= P(aX + b \leq y)$$

$$= \begin{cases} P\left(X \leq \frac{y-b}{a}\right) & a > 0 \\ P\left(X \geq \frac{y-b}{a}\right) & a < 0 \end{cases}$$

$$= \begin{cases} F_X\left(\frac{y-b}{a}\right) & a>0 \\ 1 - F_X\left(\frac{y-b}{a}\right) + P\left(X = \frac{y-b}{a}\right) & a<0 \end{cases}$$

If X is continuous RV, then

$$F_Y(y) = \begin{cases} F_X\left(\frac{y-b}{a}\right) & a>0 \\ 1 - F_X\left(\frac{y-b}{a}\right) & a<0 \end{cases}$$

$$\frac{d}{dy} F_Y(y) = f_Y(y) = \begin{cases} \frac{1}{a} f_X\left(\frac{y-b}{a}\right) & a>0 \\ -\frac{1}{a} f_X\left(\frac{y-b}{a}\right) & a<0 \end{cases}$$

Let $Y = X^2$

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

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

$= P(|X| \le \sqrt{y})$

$= F_X(\sqrt{y}) - F_X(\sqrt{y}) + P(X = -\sqrt{y}),\ y > 0$

for continuous case,

$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}),\quad y > 0$

$\Rightarrow f_Y(y) = \dfrac{f_X(\sqrt{y})}{2\sqrt{y}} + \dfrac{1}{2\sqrt{y}} f_X(-\sqrt{y}),\ y > 0$

Eg: Let $X \sim N(0,1)$. Consider the transformation $Y = X^2$

Find pdf of Y.

$\Rightarrow f_X(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2},\quad -\infty < x < \infty$

$f_Y(y) = \dfrac{1}{2\sqrt{y}} \left[\dfrac{1}{\sqrt{2\pi}} e^{-y/2} + \dfrac{1}{\sqrt{2\pi}} e^{-y/2} \right]$

$\boxed{f_Y(y) = \dfrac{1}{\sqrt{2\pi y}} e^{-y/2}}\quad 0 < y < \infty$

$y \sim G(\frac{1}{2}, 2)$

$f_X(y, \frac{1}{2}, 2)$

Eg: Let $X \sim N(0,1)$, $Y = |X|$

pdf of Y is known as half normal distribution

MGF approach:

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

$Y = n - X$

What is the prob. dist. of Y

$$M_Y(t) = E(e^{tY})$$

$$= E(e^{t(n-X)}) = e^{tn} E[e^{-Xt}]$$

$$= e^{tn} M_X(-t)$$

$$= e^{tn} (q + pe^{-t})^n$$

$$\boxed{M_Y(t) = (p + qe^{t})^n} \Rightarrow$$ this is the MGF of a $B(n,q)$ variable

$\Rightarrow$ $\boxed{Y \sim B(n,q)}$ distⁿ