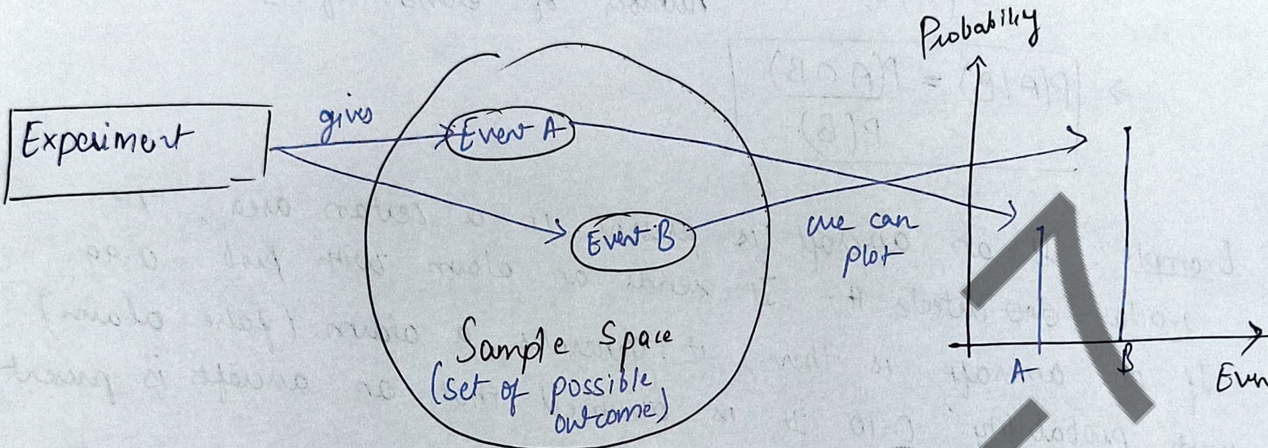
PROBABILITY



Probabilistic Model:

* Every probabilistic model involves an underlying process. This process is called Experiment.
* It produces exactly one out of several outcomes.
* Set of all possible outcome is called sample space of the experiment (Ω)
* A subset of the sample space is called event.

Definition: The probability of A satisfying the following axioms -

1) (Non-negativity) : $\boxed{P(A) \geq 0}$ for every A.
2) (Additivity) : If A & B are two disjoint events then
$$\boxed{P(A \cup B) = P(A) + P(B)}$$
3) (Normalization) : The probability of entire sample space is 1
$$\boxed{P(\Omega) = 1}$$

Properties of Probability laws: Let A, B & C be event from an experiment

(i) If $A \subset B$ then $P(A) \leq P(B)$

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

Conditional Probability: If all the outcomes are finitely many & equal likely.

$$P(A/B) = \frac{\text{number of element of } A \cap B}{\text{number of element of } B}$$

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

Example: If an aircraft is present in a certain area, the radar ~~det~~ detects it. It sends an alarm with prob 0.99. If no aircraft is there, it generates a alarm (false alarm) with probability 0.10. It is assumed that an aircraft is present with probability 0.05.

P_1(no aircraft present and a false alarm) = ?

P_2(aircraft present and no ~~direct~~ detection) = ?

Let A & B be the events.

$A = \{$an aircraft is present$\}$

$B = \{$the radar generates an alarm$\}$

$P_1(A^c \cap B) = ?$ $\quad P_2(A \cap B^c) = ?$

We know that

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

$$P_1(A^c \cap B) = P(A^c) \times P(B/A^c) = 0.95 * 0.1 = 0.095$$

$$P_2(A \cap B^c) = P(A) \times P(B^c/A) = 0.05 \times 0.01 = 0.0005$$

Multiplication Rule: Suppose an event A occurs if and only if each each of several event $A_1, A_2, \ldots, A_n$ has occured

i.e. $A = A_1 \cap A_2 \cap \ldots A_n$

Definition $P(A) = P(\bigcap_{i=1}^{n} A_i)$

$= P(A_1) \cdot P(A_2/A_1) \cdot P(A_2) \cdot P(A_3/A_2 \cap A_1)$

$$\boxed{P(A) = P(A_n / \bigcap_{i=1}^{n-1} A_i)}$$

assuming that all the conditioning events have positive probability.

The multiplication rule of probability is used to calculate the joint probability of two or more events occurring at the same time. The rule is applicable to both dependent and independent events.

Ex - Three cards are drawn from deck of 52 cards without replacement. what is the probability that none of these there card is "heart"?

Let $A_1 = \{1^{st}$ card is not a heart$\}$

$A_2 = \{2^{nd}$ card is not a heart$\}$

$A_3 = \{3^{rd}$ card is not a heart$\}$

$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_2 \cap A_1)$

Now, $P(A_1) = P(1^{st}$ card is not a heart$) = 39/52$

$P(A_2/A_1) = P(2^{nd}$ card is not a heart $/ 1^{st}$ card is not heart$) = 38/51$

$P(A_3/A_1 \cap A_2) = P(3^{rd}$ card is not a heart $/ 1^{st} \,\&\, 2^{nd}$ is not heart$) = 37/50$

$P(A_1 \cap A_2 \cap A_3) = \frac{39}{52} \times \frac{38}{51} \times \frac{37}{50} = 0.4135$

Total probability theorem: Let $A_1, A_2 \ldots A_n$ be disjoint events that form a partition of sample space. Assume that $P(A_i) > 0 \;\forall i$

Then $\boxed{P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \cdots + P(A_n) \cdot P(B/A_n)}$

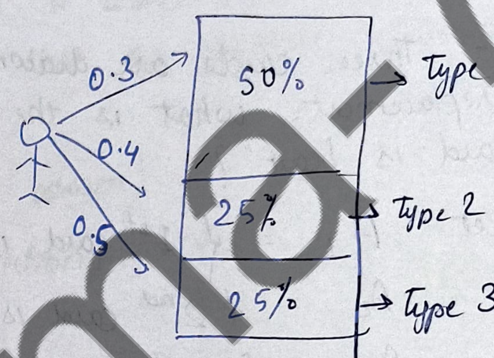
"There are number of causes that may result in certain effects. we observe the causes and wish to infer the effect"

Ex- You are playing chess tournament. Your probability of winning a game is 0.3 against half of the players. The prob of winning is 0.4 against a quarter of the players. The prob of winning is 0.5 against a quarter of the players. What is the probability that you will win?

A_1 = the event of playing with type 1.
A_2 = the event of playing with type 2.
A_3 = the event of playing with type 3

$P(A_1) = \frac{1}{2}, \; P(A_2) = \frac{1}{4}, \; P(A_3) = \frac{1}{4}$

Let B = event of winning



$P(B/A_1) = 0.3, \; P(B/A_2) = 0.4$
$P(B/A_3) = 0.5$

$P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + P(A_3) \cdot P(B/A_3)$

$= \frac{1}{2} \times 0.3 + \frac{1}{4} \times 0.4 + \frac{1}{4} \times 0.5 = 0.375$

Baye's Rule: It is used for inference. "There are a number of causes that may result in certain effect. We observe the effect & we wish to infer the cause"

Definition: Let $A_1, A_2 \ldots A_n$ be disjoint events that form a partition of the sample space & assume that $P(A_i) > 0$. Then for any event B $(P(B) > 0)$ we have:

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

Ex- Relook at the previous example of chess tournament. Suppose you win. What is the probability that you played against type 1 player?

$P(A_1/B) = \frac{P(A_1) \cdot P(B/A_1)}{P(B)} = \frac{\frac{1}{2} \times 0.3}{0.375}$

Independence: Let A & B be two events. the interesting case is, when the occurence of B provide no information & does not alter the probability of occurence of A.

$$P(A/B) = P(A) \quad \text{—— } ①$$

When Eq ① holds, then event A & B are said to be independent. from conditional prob, we know that

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

$$\Rightarrow \boxed{P(A \cap B) = P(A) \cdot P(B)} \quad \text{—— } ②$$

Eq ② is adopted as the definition of independence of two event because it can be used even when $P(B) = 0$, in which case $P(A/B)$ is undefined.

→ Independence is a ~~Syst~~ Symmetric property:
If A is independente of B, then B is independent of A.

→ If two disjoint event A & B. with $P(A) > 0$, $P(B) > 0$ &

$$A \cap B = \phi \Rightarrow P(A \cap B) = 0$$

$$\because P(A) \text{ \& } P(B) > 0 \Rightarrow P(A) \cdot P(B) > 0$$

⇒ Events are dependent

The independence of several events: The events $A_1, A_2 \dots A_n$ are said to be independent if,

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

for every subset S of $\{1, 2 \dots n\}$

Example: Let A_1, A_2, A_3 be three event. A_1, A_2, A_3 are said to be independent if and only if:

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

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$

$$P(A_2 \cap A_3) = P(A_2) \cdot P(A_3)$$

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

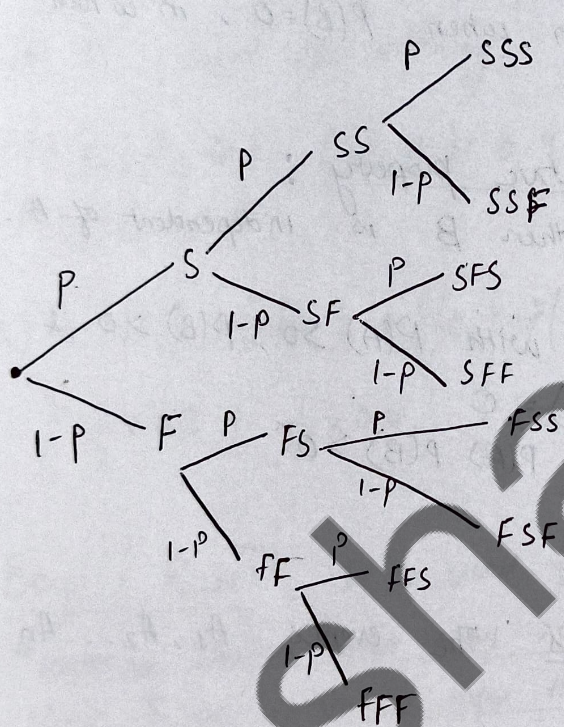
Independent trials & Binomial probabilities:

If an experiment involves a sequence of independent and identical stages. we say that we have a sequence of independent trails.

Suppose there are only two possible results at each stage, these trails are called Bernoulli trials.

Let prob of success is p & prob of failure is $1-p$.

success $\rightarrow$ S
failure $\rightarrow$ F



$P(3 \text{ success}) = p^3$

$P(2 \text{ success out of } 3 \text{ trail}) = p^2(1-p)$

$P(2 \text{ success out of } 3 \text{ trail}) = p^2(1-p)$

$P(1 \text{ success out of } 3 \text{ trail}) = p(1-p)^2$

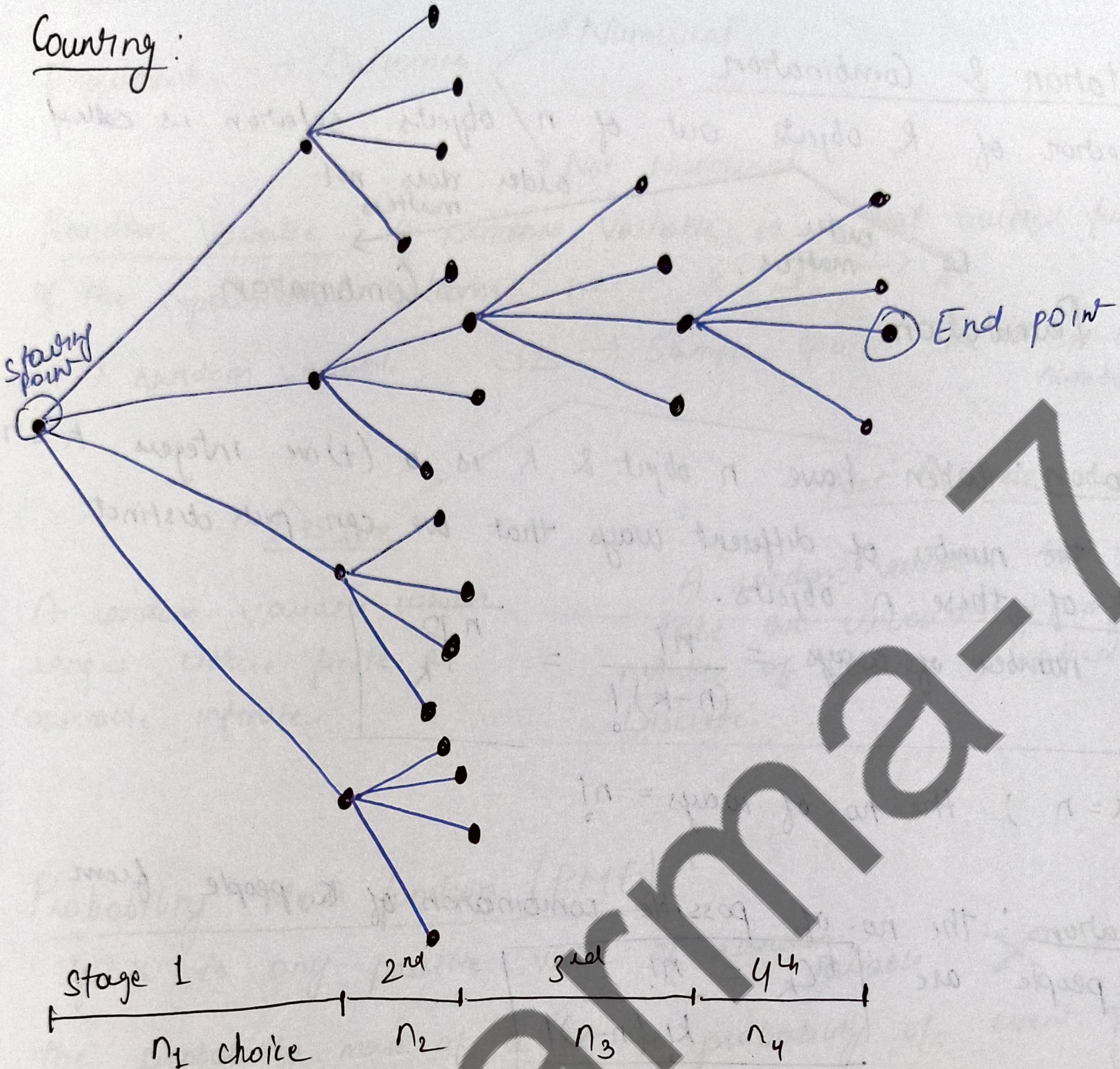
$P(0 \text{ success out of } 3 \text{ trail}) = (1-p)^3$

the probability that we get k success in n trail is

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

The number $\binom{n}{k}$ are know as Binomial Coefficients, the probabilities p^k are known as the Binomial Probabilities.

Counting :



The 1^{st} stage has n_1 possible result. for every possible result at the first $i-1$ stages. there are n_i possible at the i^{th} stage. then total numbers of possible results of the k stage is

$$\boxed{n_1 . n_2 \ldots n_k}$$

Example: A telephone number is a 7 digit sequence, but the 1^{st} digit should not be one or zero. How many distinct telephone numbers are there.

8	10	10	10	10	10	10
1 no	2^{nd}	3^{rd}	4^{th}	5	6	7
8 choices	10	10	10	10	10	10

The answer $= 8 \times 10^6$.

Permutation & Combination :

Selection of R objects out of n objects. selection is called

order matters → Permutation.

order does not matters → Combination

Permutation : when have n object & k is a (+)ive integer $k \leq n$

Count the number of different ways that we can pick distinct k out of these n objects.

$$\boxed{\text{the number of ways} = \frac{n!}{(n-k)!} = {}^{n}P_{k}}$$

$\because k = n$; the no. of ways $= n!$

Combination : The no. of possible combination of k people from n people are $\boxed{{}^{n}C_{k} = \frac{n!}{k!\,(n-k)!}}$

for ex: we have 4 English alphabet i.e. A, B, C & D.

Permutation : AB, BA, BC, CB, AD, DA, BD, DB, CD, DC, AC, CA.

Combination : AB, AC, AD, BC, BD, CD.

Experiment → Outcomes → Numerical
Outcomes → Not Numerical

Random Variable: A random Variable is a real valued function of the experimental outcomes i.e. $X : \Omega \rightarrow \mathbb{R}$

$X \rightarrow$ random variable, $\Omega \rightarrow$ Sample space, $\mathbb{R} \rightarrow$ Set of real number.

Discrete

A random variable which ranges either finite or countable infinite.

Not Discrete

A random variable that can take an uncountably infinite number of values is called Not Discrete.

Probability Mass function (PMF):

If x is any possible value of random variable X, the probability mass of x is the probability of event $\{X = x\}$

i.e. Probability mass of $x = p_X(x) = P\{X = x\}$

Example: Experiment: tossing a fair coin twice. random variable

X : no. of obtained $X = 0, 1, 2.$ $x = 0, 1, 2$

PMF of random Variable X

$$p_X(x) = \begin{cases} 1/4 ; & x = 0 \\ 1/2 ; & x = 1 \\ 1/4 ; & x = 2 \end{cases}$$

NOTE: $\boxed{\sum_x p_X(x) = 1}$ sum of all event is 1.

Bernoulli Distribution (fail or success):

X is said to be Bernoulli random variable

If $X = \begin{cases} 1 ; & \text{if success} \\ 0 ; & \text{if failure} \end{cases}$

Suppose probability of success is p. then its PMF is

$$p_X(x) = \begin{cases} p & ; \text{ if } K = 1 \\ 1-p & ; K = 0 \end{cases}$$

$\boxed{\text{mean} = p}$ $\boxed{\text{Varience} = p(1-p)}$

Example: ① A telephone at a given time can be either free or busy.

② A person can be healthy or sick.

X

Binomial Random Variable:

Let X be the number of success in a n-trail. X is called Binomial Random Variable with parameters n & p. the PMF

$$\boxed{p_X(x) = p(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2 \dots n}$$

$\boxed{\text{mean} = np}$ $\boxed{\text{Varience} = np(1-p)}$

Example: ① Number of defective item in a batch.

② Coin flips resulting in head.

The normalization property give:

$$\boxed{\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1}$$

Geometric Random Variable: The geometric random variable is the number X of trail to get the first sucass. Its PMF is given by.

$X = 1, 2, 3 \ldots$

Suppose prob of sucass is p

$$\boxed{P(X=k) = p_X(k) = (1-p)^{k-1} p, \quad k = 1, 2, 3 \ldots}$$

$$\boxed{\text{mean} = 1/p} \qquad \boxed{\text{Varience} = \frac{1-p}{p^2}}$$

$$\sum_{k=1}^{\infty} p_X(k) = 1$$

Proof: $\sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p = p \cdot \sum_{k=1}^{\infty} (1-p)^{k-1}$

$= p[(1-p)^0 + (1-p)^1 + (1-p)^2 \ldots]$

$= p[1 + (1-p) + (1-p)^2 \ldots]$

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

The Poisson Random Variable: It models the numbes of events occuring in a fixed interval of time or space, when events occur independently and at a constant average rate. PMF is given by

$$\boxed{p_X(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2 \ldots \; \& \; \lambda > 0}$$

$$\boxed{\text{mean} = \lambda} \qquad \boxed{\text{Varience} = \lambda}$$

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = 1$$

Relation b/w Poisson & Binomial:

Take a binomial random variable with very small p and very large n.

The Poisson PMF with parameter λ is a good approximation for a binomail PMF with parameter n & p, i.e.

$$\frac{e^{-\lambda}\lambda^k}{k!} \approx \binom{n}{k} p^k (1-p)^{n-k} \quad \text{—— ①}$$

$k = 0, 1, \ldots n$

provided $\lambda = np$

for example: Let X be the number of typos in a book with a total of n words. Then X is binomial random variable.

$n = 100$, $p = 0.01$, then probability of 5 success in 100 trail is.

binomail $\rightarrow \binom{100}{5} (0.01)^5 (1-0.01)^{100-5} = 0.0029$

$\lambda = np = 100 \times 0.001 = 1$

Using Poisson PMF with

$$P(X=5) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-1}}{5!} = 0.0030$$

Function of random Variable: Let X be a random variable.

If $Y = g(X)$ is a function of X, then Y is also a random variable. Since, it provide a numerical value for each possible outcome.

If X is discrete then Y is also discrete random variable, the PMF of Y is calculated from PMF of X.

the PMF of Y is

$$\boxed{p_Y(y) = \sum_{\{x \mid g(x) = y\}} p_X(x)}$$

Example: Let X be a random variable with PMF

$$p_X(x) = \begin{cases} 1/9 & \text{if } x \text{ is a integer in } [-4, 4] \\ 0 & ; \text{ otherwise} \end{cases}$$

Let $Y = |X| \Rightarrow Y$ is a discrete random variable

Possible value of $Y = 0, 1, 2, 3, 4$

$$P(Y=0) = P(X=0) = 1/9$$
$$P(Y=1) = P(X=1) + P(X=-1) = 2/9$$
$$P(Y=2) = P(X=2) + P(X=-2) = 2/9$$
$$P(Y=3) = P(X=3) + P(X=-3) = 2/9$$
$$P(Y=4) = P(X=4) + P(X=-4) = 2/9$$

$$p_Y(y) = \begin{cases} 1/9 & ; y = 0 \\ 2/9 & ; y = 1, 2, 3, 4 \\ 0 & ; \text{otherwise} \end{cases}$$

Expectations of a random variable:

We define the expected value of a random variable as follows:

$$\boxed{E[X] = \sum_x x \cdot p_X(x)}$$ where p_X is PMF of X

Example: Let X be a random variable with PMF

$$\text{PMF } p_X(x) = \begin{cases} 1/9 & ; x \text{ is an integer } \& \ x \in [-4, 4] \\ 0 & ; \text{otherwise} \end{cases}$$

find $E[X]$?

$$E[X] = \left(-4 \times \frac{1}{9}\right) + \left(-3 \times \frac{1}{9}\right) + \left(-2 \times \frac{1}{9}\right) + \left(-1 \times \frac{1}{9}\right) +$$
$$\left(0 \times \frac{1}{9}\right) + \left(1 \times \frac{1}{9}\right) + \left(2 \times \frac{1}{9}\right) + \left(3 \times \frac{1}{9}\right) + \left(4 \times \frac{1}{9}\right) = 0$$

Q. What is difference b/w average & p. expectations?
In average all events have same prob where as in expectation the prob are different because we have random variable.

Expectation of function of random Variable:

Let x be a random variable with PMF p_x & Let $g(x)$ be the function of x. then

$$\boxed{E[g(x)] = \sum_x g(x)\, p_x(x)} \quad (1)$$

Property of Expectation:

Let x be a random variable & let

$$\boxed{Y = aX + b} = g(x) \quad \left(\text{a and b are scalars}\right)$$

$$E[Y] = E[g[x]] = \sum_x (ax+b)\, p_x(x)$$

$$= a\sum_x x\, p_x(x) + b\sum_x p_x(x) = aE[X] + b$$

$$\Rightarrow \boxed{E[Y] = aE[X] + b}$$

Varience: Varience of random variable X is defined as.

$$\boxed{Var(X) = E[X - E[X]]^2}$$

or

$$\boxed{Var(x) = E[x^2] - (E[x])^2}$$

Example: Let X be a random Variable with PMF :

$$p_X(x) = \begin{cases} 1/9 & ; x \text{ is integer } \& \ x \in [-4,4] \\ 0 & ; \text{otherwise} \end{cases} \quad Var(X) = ?$$

we have already calculate $E[X] = 0$

so, $Var[X] = E[X - E[X]]^2 = E[X-0]^2 = E[X^2]$

so, we need PMF X^2

$$p_Z(z) = \begin{cases} 1/9 & z=0 \\ 2/9 & z = 1,4,9,16 \\ 0 & \text{otherwise} \end{cases}$$

$$E[Z] = 0 \times \frac{1}{9} + 1 \times \frac{2}{9} + 4 \times \frac{2}{9} + 9 \times \frac{2}{9} + 16 \times \frac{2}{9} = \frac{60}{9}$$

$Var[X] = 60/9$

Standard deviation: the underoot of varience, deviation from the mean. denoted by σ_X

$$\boxed{\sigma_X = \sqrt{Var(X)}}$$

Property of Varience:

① Varience is always non-negative

② Let X be a random variable & Let

$$\boxed{Y = aX + b} \quad (a \ \& \ b \text{ are scalars})$$

$$Var[g(x)] = E[g(x) - E[g(x)]]^2$$

$$= \sum_x \left(ax + b - E[ax + b]\right)^2 p_X(x)$$

$$= a^2 \sum (x - E[x])^2 p_X(x) = a^2 E[(X - E[X])^2]$$

$$\boxed{Var(Y) = a^2 Var(X)}$$

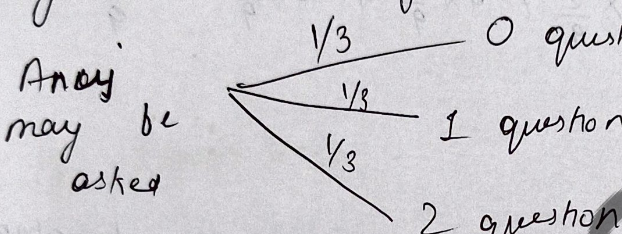
* Joint PMF of Multiple random Variable :

Consider two discrete random variable X & Y associated with same experiment.

Then joint PMF of X & Y is :-

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

Example: Prof Anuj answer each of his students question incorrectly with probability $\frac{1}{4}$ independent of other question.



X = number of question he is asked

Y = no. of question he answered incorrectly.

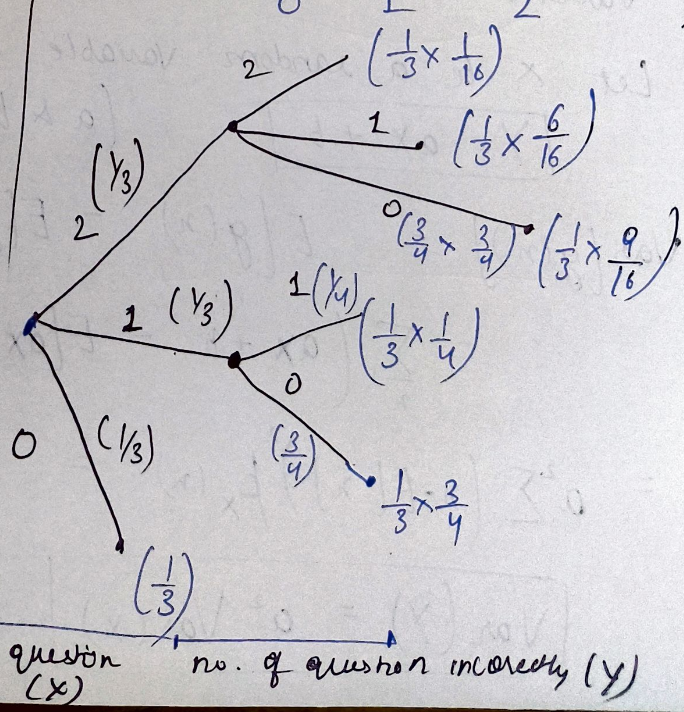
Joint PMF of X & Y

$Y \backslash X$	0	1	2
2	0	0	1/48
1	0	1/12	2/16
0	1/3	1/4	3/16

$$p_{X,Y}(1,1) = P(X=1, Y=1)$$

$$= p_{Y|X}(y|x)\, p_X(x)$$

$$= \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$$



NOTE: $E[a_1 X_1 + a_2 X_2 + \dots + a_n X_n] = a_1 E[X_1] + a_2 E[X_2] + \dots + a_n E[X_n]$

where $a_1, a_2, a_3 \dots a_n$ are constant.

Conditioning: Let X & Y be two random Variables associated with same experiment. the conditional probability

$$p_{X|Y}(x/y) = P(X=x / Y=y)$$

from conditional probability definition

$$p_{X|Y}(x/y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{p_{XY}(x,y)}{p_Y(y)}$$

$$\boxed{p_{X/Y}(x/y) = \frac{p_{XY}(x,y)}{p_Y(y)}}$$

Example: Extended ~~version~~ question of Prof Anuj.

① P(at least one wrong answer)

$$P(X=1, Y=1) + P(X=2, Y=1) + P(X=2, Y=2)$$

$$= \frac{1}{12} + \frac{2}{16} + \frac{1}{48} = \frac{11}{48}$$

Independence of random Variable:

Two random Variable are independent if

$$\boxed{p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y) \quad \forall x, y}$$

* $E[X \cdot Y] = E[X] \cdot E[Y]$
* $Var(X_1 + X_2) = Var(X_1) + Var(X_2)$

Continous random Variables:

A random Variable X is called continous if there is non-negative function f_x such that

$$\boxed{P(X \in B) = \int_B F_X(x)\, dx}$$

[where B is a subset of real line]

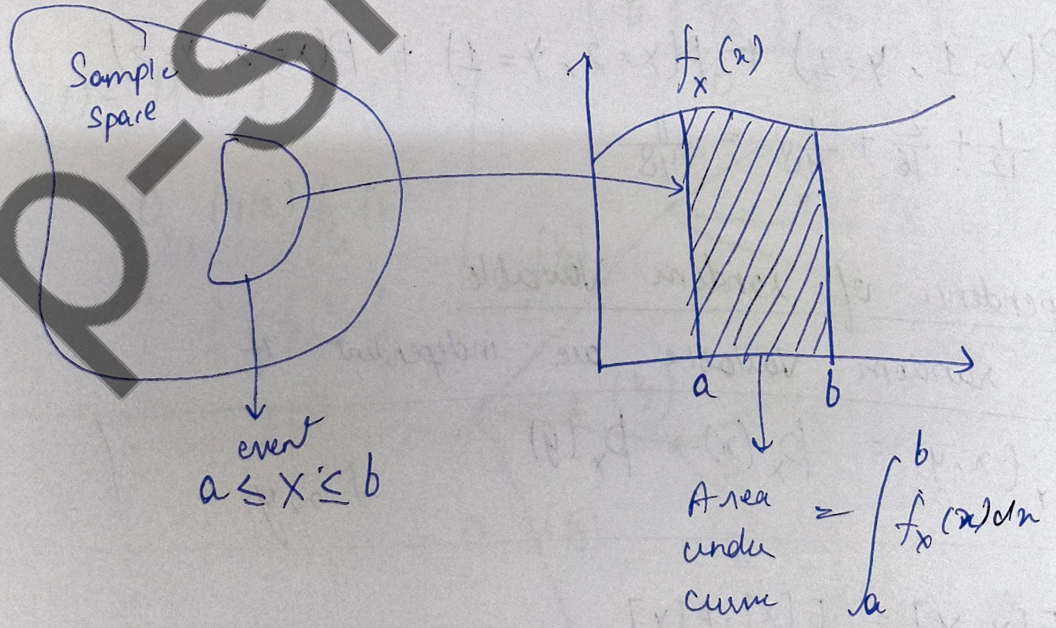
or

$$\boxed{P(a \le X \le b) = \int_a^b F_X(x)\, dx}$$

the function $f_X(x)$ is called the probability density function (PDF) of X.

For Example: $p(1 \le X \le 2) = \int_1^2 f_X(x)\, dx$

$p(2 \le X \le 3) = \int_2^3 f_X(x)\, dx$



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

Uniform random Variable :

Consider a random variable X that takes values in an interval $[a, b]$ assuming that any two sub interval of same length have same probability. such random variable are called uniform random variable.

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} ; & a \leq x \leq b \\ 0 ; & \text{otherwise} \end{cases}$$

$$\circledast \int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{a} 0\,dx + \int_{a}^{b} \frac{1}{b-a}\,dx + \int_{b}^{\infty} 0\,dx$$

$$= \frac{1}{b-a}[b-a] = 1$$

A PDF can take arbitary large value :

Let X be a random variable with PDF

$$f_X(x) = \begin{cases} \dfrac{1}{2\sqrt{x}} ; & 0 \leq x \leq 1 \\ 0 ; & \text{otherwise} \end{cases}$$

$$\boxed{\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{0}^{1} \frac{1}{2\sqrt{x}}\,dx = 1}$$

NOTE :
1. PDF can never be negative
2. PDF can be infinitely large.
3. PDF can be any value but normalization of pdf be equal to 1.

Expectation: The expectation of a continous random variable X is defined as :

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

If $g(X)$ is a function of n random variable X, then

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

for ex: $g(X) = X^2$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx \quad \text{—— (2)}$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 \quad \text{—— (3)}$$

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

Mean & Varience of Uniform Random :

Let X be a uniform random variable then its PDF is

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

$$E[X] = \int_a^b x\left(\frac{1}{b-a}\right) dx = \left(\frac{1}{b-a}\right)\left[\frac{x^2}{2}\right]_a^b = \frac{a+b}{2}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \left(\frac{1}{b-a}\right)\left[\frac{x^3}{3}\right]_a^b = \frac{a^2+b^2+ab}{3}$$

$$Var(x) = E[x^2] - (E[x])^2 = \left(\frac{a^2+b^2+ab}{3}\right) - \left(\frac{a+b}{2}\right)^2$$

$$\boxed{Var(x) = \frac{(b-a)^2}{12}} \qquad \boxed{E[x] = \frac{a+b}{2}}$$

Exponential Random Variable

An Exponential random variable is used to model the amount of time until an incident of interest takes place.

for example:

① Breaking down of some equipment

② Burning out of a light bulb.

③ An accident occuring.

If X is an exponential random variable,

$$P(X \geq a) = \int_a^{\infty} f_x(x)\,dx = \int_a^{\infty} \lambda e^{-\lambda x}\,dx$$

$$\boxed{P(X \geq a) = e^{-\lambda a}}$$

$$\boxed{E[x] = \frac{1}{\lambda}} \qquad \boxed{Var[x] = \frac{1}{\lambda^2}}$$

Ex: The time until a small meteorite first lands anywhere in Sahara desert is modelled as an exponential rv with mean of 10 days the time is currently mid-night What is the prob that a meteoric first lands some time b/w 6 am to 6 pm of the first day?

Let X be the time elapsed until the event of interest. It is given that.

$$E[X] = 10$$

we know that $E[X] = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{10}$

$$P\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right) = P\left(X \geq \frac{1}{4}\right) - P\left(X \geq 3/4\right)$$

$$= e^{-\frac{1}{40}} - e^{-3/40} = 0.0476$$

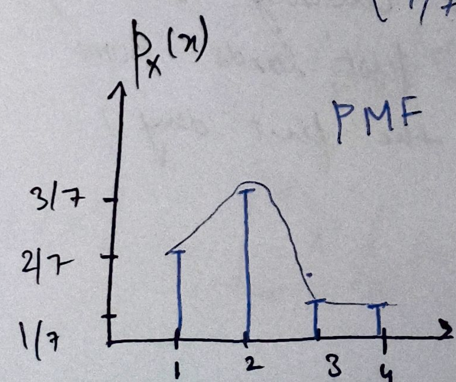
Cumulative Distribution function (CDF)

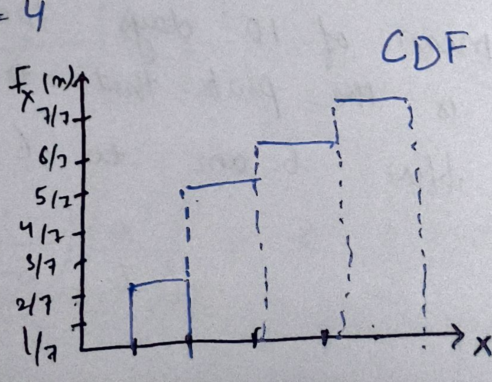
The CDF of a random variable X is denoted by F_X & provide the probability $P(X \leq x)$. for every x, we have,

$$F_X(x) = P(X \leq x) = \begin{cases} \sum_{k \leq x} p_X(x) & ; \text{ If } X \text{ is discrete} \\ \int_{-\infty}^{k} f_X(x)\,dx & ; \text{ If } X \text{ is continous} \end{cases}$$

Example: Let X be a discrete random variable

$$p_X(x) = \begin{cases} 2/7 & ; x = 1 \\ 3/7 & ; x = 2 \\ 1/7 & ; x = 3 \\ 1/7 & ; x = 4 \end{cases}$$

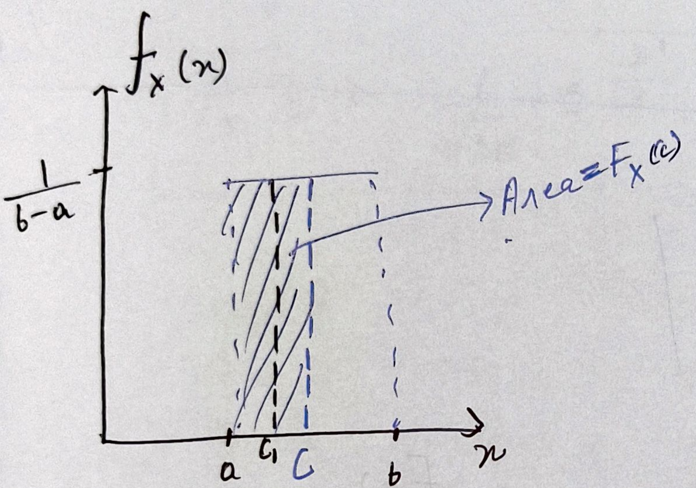




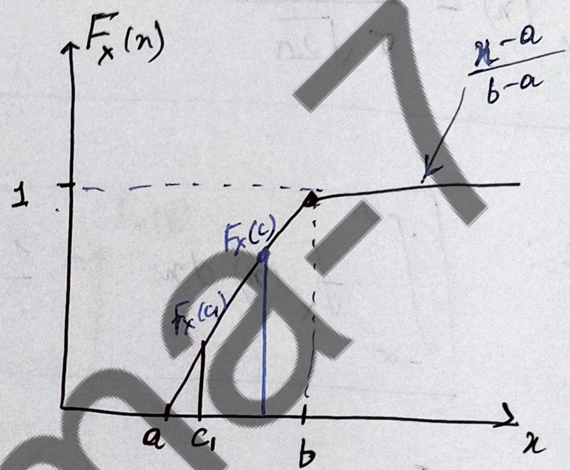
~~Let~~ ~~&~~

Let X be a continous (Uniform RV) :

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



PDF



CDF

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{b-a}\, dt = \int_{a}^{x} \frac{1}{b-a}\, dt = \frac{x-a}{b-a}$$

Properties of CDF :

1. If $x \le y$ then $F_X(x) \le F_Y(y)$
2. $F_X(x)$ tends to as $x \longrightarrow -\infty$ and tends to 1 as $x \longrightarrow \infty$
3. If X is discrete then $F_X(x)$ is a piecewise constant fn.
4. If X is continous then $F_X(x)$ is a continous function.
5. If X is continous then

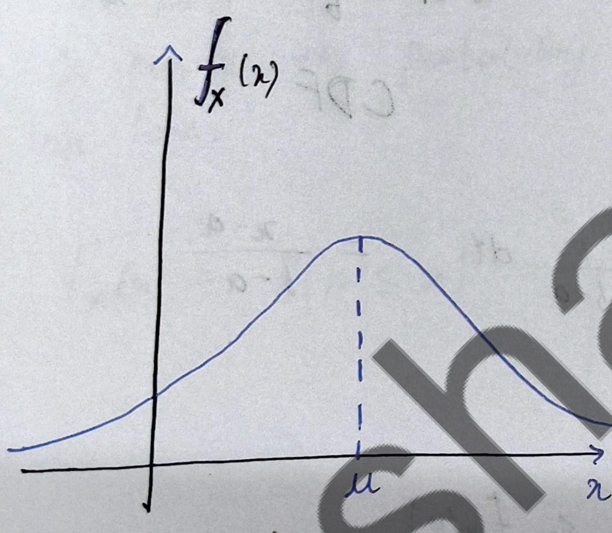
$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt \quad \& \quad f_X(t) = \frac{dF_X(x)}{dx}$$

Normal random Variable :

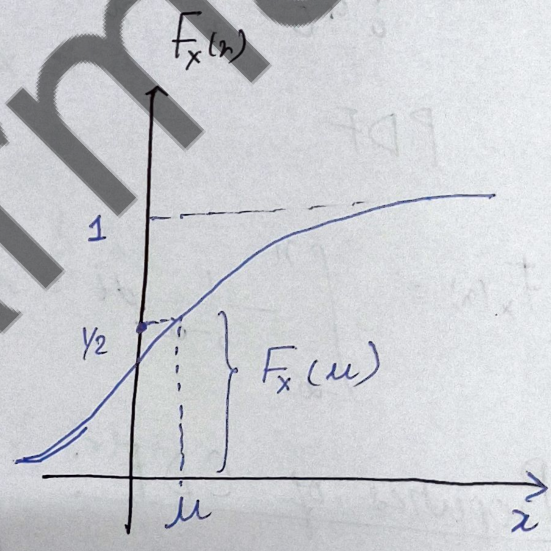
A continuous random variable X is said to be normal (or Gaussian) if it has a PDF of the form.

$$\boxed{f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}} \quad \text{where } -\infty < x < \infty}$$

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



PDF



CDF

The mean & Varience :

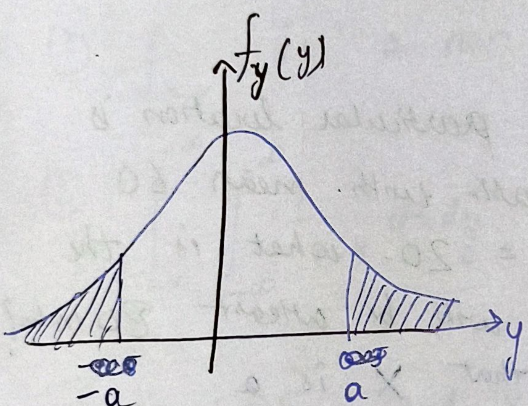
$$E[X] = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\}} dx = \mu$$

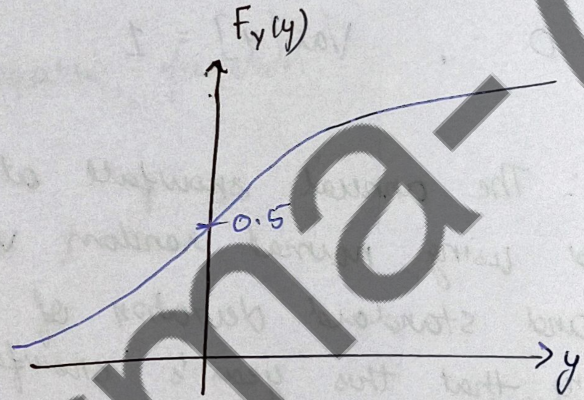
$$Var[X] = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\}} dx = \sigma^2$$

Standard normal random variable:

A normal random variable with $\mu = 0$ & $\sigma^2 = 1$ is said to be standard normal random variable (denoted by Y). It's PDF will be

$$\boxed{f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}}$$





$$\Phi(-a) = P(Y \le -a)$$
$$= 1 - P(Y \le a)$$
$$= 1 - \Phi(a)$$

In general

$$\boxed{\Phi(-y) = 1 - \Phi(y)}$$

Mean and Variance:

Let X be a normal random with mean μ and variance σ^2

Define random variable. $Y = \frac{X - \mu}{\sigma}$

Y is a linear function, Hence it is normal random variable.

$$E[Y] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{E[X]}{\sigma} - \frac{\mu}{\sigma} = \frac{\mu-\mu}{\sigma} = 0$$

Of

$$Var[Y] = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(X) \quad (2)$$

$$Var[Y] = \frac{1}{\sigma^2} \times \sigma^2 = 1.$$

$$E[Y] = 0 \quad , \quad Var[Y] = 1$$

<u>Example</u> : The annual snowfall at a particular location is modeled using normal random variable with mean 60 inches and standard deviation of $\sigma = 20$. What is the probability that this year's snowfall will be atleast 80 inches?

Let X be the snowfall. It is given that X is a normal random variable with $\mu = 60$ & $\sigma = 20$

$$Y = \frac{X-60}{20} \qquad \boxed{Y = \frac{X-\mu}{\sigma}}$$

We know that Y is a standard normal random variable.

We need to find

$$P(X \geq 80) = P\left(\frac{X-60}{20} > \frac{80-60}{20}\right) = P(Y \geq 1)$$

$$= 1 - P(Y \leq 1) = 1 - 0.8413$$

Joint PDF of Multiple random Variable:

Let X & Y be two continous random variable associated with same experiment the joint PDF of X & Y will be.

$$\boxed{P(a \le X \le b,\ c \le Y \le d) = \int_c^d \int_a^b f_{(X,Y)}\, dx\, dy}$$

① $f_{X,Y}$ is a non negative function

② $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}\, dx\, dy = 1$

Joint CDF: Let X & Y be two random variable associated with same experiment, their joint CDF is given by

$$\boxed{F_{X,Y}(x,y) = P(X \le x,\ Y \le y)}$$

If X & Y are continous then

$$\boxed{F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y)}$$

we have

$$\boxed{\frac{d^2 F_{X,Y}(x,y)}{\partial x \partial y} = f_{X,Y}(x,y)}$$

Expectation: If X & Y continous RV.

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Conditioning one random variable on another:

Let X & Y be continous random variable with joint PDF f_{XY}: The conditional PDF of X given $Y = y$ is defined as.

$$\boxed{f_{X/Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}}$$

Ex: The speed of a typical vehicle that drives past ~~a place~~ a police radar is modeled as an exponential random varible X with mean 50 miles/hr. The police radar's measurement Y of the vehicles's speed radar's measurement Y of the vehicles speed has an error which modelled as a normal random variable with zero mean & standard deviation equal to one-tenth of the vehicle's speed what is the Joint PDF of X & Y.

X is the exponential random variable & $\lambda = \frac{1}{50}$ $\left[\because \text{mean} = 50 = \frac{1}{\lambda}\right]$

Given that $X = x$, the measurement Y has a normal PDF with mean x & variance $= \frac{x^2}{100}$

$$\boxed{f_{Y/X}(y/x) = \frac{1}{\sqrt{2\pi}\left(\frac{x}{10}\right)} e^{\left(\frac{-y+x}{2\left(\frac{x}{10}\right)^2}\right)}}$$

$$f_{Y/X}(x,y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

$$f_{XY}(x,y) = f_{Y/X}(x,y) \cdot f_X(x)$$

$$= \lambda e^{-\lambda x} \cdot \frac{1}{\sqrt{2\pi}\left(\frac{x}{10}\right)} e^{\frac{-(y-x)^2}{2x^2/100}}$$

$$= \frac{1}{50} e^{-x/50} \cdot \frac{1}{\sqrt{2\pi}\left(\frac{x}{10}\right)} e^{-50}$$

<u>Independence</u>: Two continuous random variable X & Y are independent if their joint PDF is the product of marginal PDFs.

$$\boxed{f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)}$$

Note that if random variable are independent then

$$\boxed{f_{X/Y}(x/y) = f_X(x)}$$

In general, if we have 3 random variable X, Y & Z these are independent if

$$\boxed{f_{X,Y,Z}(x,y,z) = f_X(x) \cdot f_Y(y) \cdot f_Z(z)}$$

Relation b/w independence & CDF

Two rv X & Y are independent then

$$\boxed{F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)}$$

Relation b/w expectation & Indenpendence :

If X & Y are independent, then

$$\boxed{E[XY] = E[X] \cdot E[Y]}$$

Relation b/w Varience & Independence

If X & Y are independent then

$$\boxed{Var(X+Y) = Var(X) + Var(Y)}$$

Derived distribution: Calculating the PDF of $Y = g(x)$ where X is a continuous random variable.

Ex - let X be uniform on $[0, 1]$ & Let $Y = \sqrt{X}$. Find PDF of Y.

PDF of X : $f_X(x) = \begin{cases} 1 & ; 0 \le x \le 1 \\ 0 & ; \text{otherwise} \end{cases}$

If x's range is $[0, 1]$ then Y's range is $[0, 1]$

$$F_x(y) = P(Y \le y) = P(\sqrt{x} \le y) = P(X \le y^2)$$

$$= \int_0^{y^2} f_X(x)\,dx = y^2 \Rightarrow \boxed{F_Y(y) = y^2}$$

then PDF of Y

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = \frac{\partial (y^2)}{\partial y} = 2y \quad ; \; 0 \le y \le 1$$

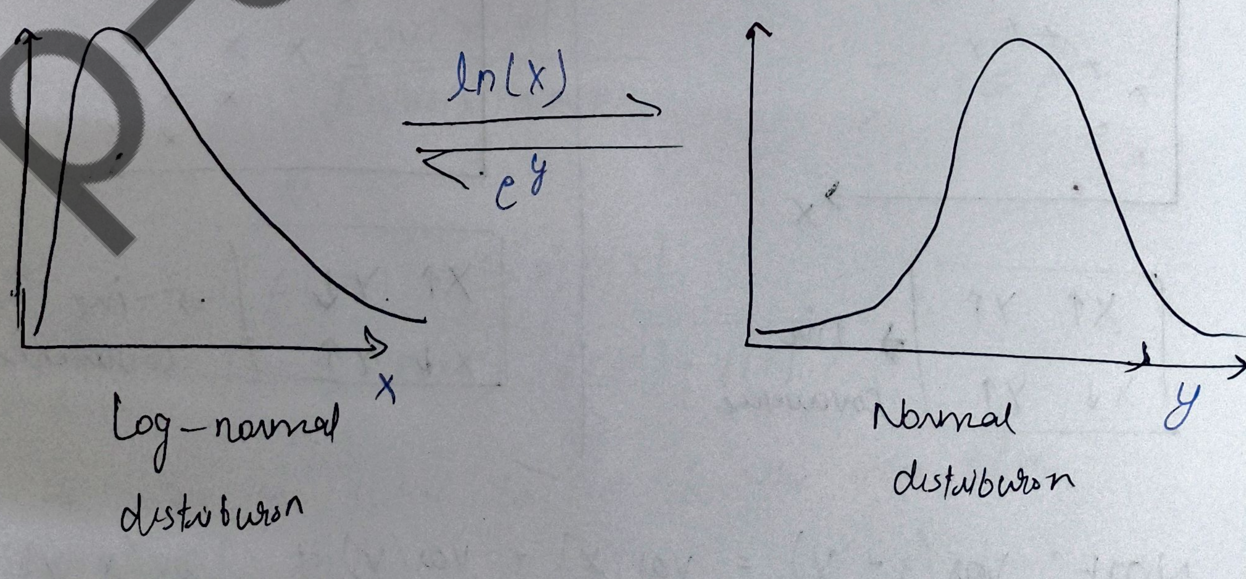
$$f_Y(y) = \begin{cases} 2y & ; \; 0 \le y \le 1 \\ 0 & ; \; \text{otherwise} \end{cases}$$

Log Normal distribution:

A postive continous random variable X is log normally distributed, if the natural logarithm of X is normally distributed with mean μ & varience σ^2.

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

$$f_X(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} e^{\left(\frac{-(\ln x - \mu)^2}{2\sigma^2}\right)} & ; \; x \ge 0 \\ 0 & ; \; \text{otherwise} \end{cases}$$



Covarience & Correlation:

① Covarience & Corelation are two statistical measures used to determine the relationship between two variables.

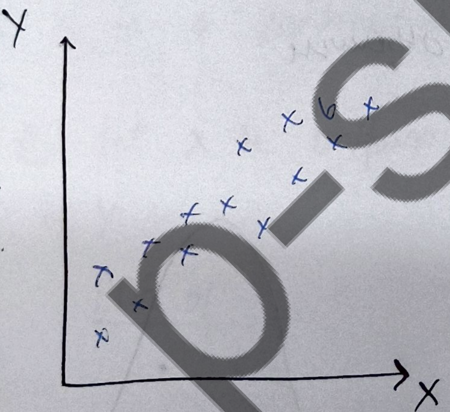
② Let x & y be two random variable. The covarience of x & y denoted by Cov(x, y) is defined as

$$Cov(x, y) = E[(x - E[x])(y - E[y])]$$

③ It is quantitive measure of the strength & direction of the relationship b/w two random variable.

Covarience: is a measure of how much two random variable change together.

① If variable tends to increase & decrease together, then the covarience is positive



X↑	Y↑
X↓	Y↑

⇒ +ive covarience

② If variable tends to decrease when the others decrease and vice & versa. the covarience is negative.

X↑	Y↓
X↓	Y↑

⇒ -ive covarience

NOTE: $Var(x + y) = var(x) + var(y) + 2cov(x, y)$

Cov (X, Y) is defined as —

$$Cov(X,Y) = \boxed{E\left[(X - E[X])(Y - E[Y])\right]}$$

OR

$$\boxed{E[XY] - E[X].E[Y]}$$

① When $Cov(X,Y) = 0$, we say that X & Y are uncorrelated.

② $Cov(X,X) = E[X^2] - E[X]E[X] = Var(X)$

③ $Cov(X, aY + b) = a\,Cov(X,Y)$

④ If X & Y are independent then $Cov(X,Y) = 0$

⑤ $Cov(X, Y+Z) = Cov(X+Y) + Cov(X,Z)$

② Correlation — Pearson correlation coefficient / Spearson correlation coefficient.

i) Pearson Correlation Coefficient: It defined as —

$$\rho_{(X,Y)} = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}} \quad ; \quad -1 \le \rho \le 1$$

as

$$\rho_{(X,Y)} = \frac{E(XY) - E[X]E[Y]}{\sqrt{(E[X^2] - E[X]^2)\,(E[Y^2] - E[Y]^2)}}$$

* the more the value toward +1, the more +ive correlated X & Y.

* The more the value toward -1, the more $-ve$ correlated it is (X, Y)

(ii) Spearman Rank Correlation:

Pearson correlation is not able capture non-linear data but Spearson rank correlation can do this.

The spearson correlation coefficient is defined as the Pearson correlation coefficient between the rank variable.

$$\boxed{\gamma_s = \rho\left[R[X], R[Y]\right] = \frac{\operatorname{cov}\left[R[X], R[Y]\right]}{\sigma_R[x]\ \sigma_R[Y]}}$$

where

ρ = It denotes conventional pearson coefficient operator but applied to the rank variable.

$\sigma_R[x]$, $\sigma_R[Y]$ = the standard deviation of rank variable

$R[X_i], R[Y_i]$ = ~~than~~ the n pair of raw score $[X_i; Y_i]$ converted to $R[X_i], R[Y_i]$

What are moment (mathematics)?

In mathematics, the moment of function are certain quantitive measures related to the shape of the function 's graph. If the function represents. mass density, then

① The zeroth moment is the total mass

② the first moment (normalized by total mass) is the centre of mass

③ the second moment is the moment of intertia.

⇒ If the function is the probability distribution then,

① the first moment is expected value.

② the second central moment is varience.

③ the third standardized moment is the skewness.

Moment generating function:

Let X be a random variable, then n^{th} moment of X is defined as $E[X^n]$.

The moment generating function associated with random variable X is defined as.

$$\boxed{M_X(s) = E[e^{sX}]}$$

where s is a scalar parameter.

for discrete random variable X:

$$\boxed{M(s) = \sum e^{sx} p_X(x)}$$

for continous random variable X:

$$M(s) = \int^{\infty} e^{sx} f_X(x)\,dx$$

Example: MGF of Poission random variable:

$$p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, 3 \dots$$

$$\text{MGF of } X = M(s) = \sum_{x=0}^{\infty} e^{sx} p_X(x) = \sum_{x=0}^{\infty} e^{sx}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{e^{sx}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^s\lambda)^x}{x!} \quad \left[\text{Let } a = e^s\lambda\right]$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{a^x}{x!} = e^{-\lambda}e^{a} = e^{a-\lambda} = e^{\lambda(e^s-1)}$$

MGF of Exponential random variable:

$$f_X(x) = \lambda e^{-\lambda x}; \quad x \geq 0$$

$$M(s) = \int_0^{\infty} e^{sx}\lambda e^{-\lambda x}\,dx = \lambda\int_0^{\infty} e^{sx}e^{-\lambda x}\,dx$$

$$= \lambda\int_0^{\infty} e^{(s-\lambda)x}\,dx = \lambda\left[\frac{e^{(s-\lambda)x}}{(s-\lambda)}\right]_0^{\infty} =$$

If $s < \lambda$:

$$\lambda\left[0 - \frac{1}{s-\lambda}\right] = \frac{\lambda}{\lambda - s}$$

Otherwise: Integral will be infinite.

from MGF to moments:

Let X be a continuous random variable then MGF of X is defined as:

$$M(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx$$

Differentiate both side with s

$$\frac{\partial M(s)}{\partial s} = \frac{\partial}{\partial s}\int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial s} e^{sx} f_X(x)\, dx$$

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

Consider the special case when $s=0$, we get

$$\boxed{\left.\frac{\partial M(s)}{\partial s}\right|_{s=0} = \int_{-\infty}^{\infty} x f_X(x)\, dx} \Rightarrow \boxed{\left.\frac{\partial M(s)}{\partial s}\right|_{s=0} = E[X]}$$

In general:

$$\boxed{\left.\frac{\partial^n M(s)}{\partial s^n}\right|_{s=0} = \int_{-\infty}^{\infty} x^n e^{sx} f_X(x)\, dx = E[X^n]}$$

Summary: Random variable X $\longrightarrow$ MGF of X

$\downarrow$

$E[X^n]$ $\longleftarrow$ Differentiate and put $s=0$

n^{th} moment

Markov Inequality: If a random variable X can take only non- ~~integer~~ negative value then

$$\boxed{P(X \geq a) \leq \frac{E[X]}{a} \quad \forall\, a > 0}$$

Chebyshev Inequality: If X is a random variable with mean μ & varience σ^2 then

$$\boxed{P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2} \quad \forall\, c > 0}$$

Ex: Let X be uniform random variable in the interval $[0, 4]$

$$f_X(x) = \begin{cases} \frac{1}{4}, & 0 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

$$E[X] = \int_0^4 x f_X(x)\,dx = \frac{1}{4}\int_0^4 x\,dx = \frac{1}{4}\left[\frac{x^2}{2}\right]_0^4 = 2$$

$$Var[X] = \frac{(b-a)^2}{12} = \frac{16}{12} = 4/3$$

Recall the markov inequality: $P(X \geq a) \leq \frac{E[X]}{a}$

for $a = 2$: $P(X \geq 2) \leq \frac{2}{2} = 1$ | $P(X \geq 2) = \int_2^4 \frac{1}{4}\,dx = 0.5$

for $a = 3$: $P(X \geq 3) \leq \frac{2}{3} = 0.67$ | $P(X \geq 3) = \int_3^4 \frac{1}{3}\,dx = 0.25$

Using Chebyshev Inequality:

$$P(|X-2| \geq 1) \leq \frac{\sigma^2}{c^2} = \frac{4}{3}$$

Imp

The Weak law of large numbers:

It states that ~~ae~~ the average of large number of independent and identically distributed (iid) observation will be close to the expected value.

If $X_1, X_2, \ldots X_n$ be a iid with mean μ. ~~for every~~

for every: $\epsilon > 0$, we have
(epsilon)

$$P\left(\left|\frac{X_1 + X_2 \ldots + X_n}{n} - \mu\right| \geq \epsilon\right) \longrightarrow 0 \quad \text{as } n \to \infty$$

Explanation: Let $X_1, X_2 \ldots X_n$ be iid with mean μ & variance σ^2. Define new ~~var~~ random variable

$$M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

$$E[M_n] = \frac{1}{n}\left[E[X_1] + E[X_2] \ldots + E[X_n]\right]$$

$$= \frac{1}{n}\left[\mu + \mu + \ldots + \mu\right] = \mu$$

$$Var[M_n] = \frac{1}{n^2}\left[\sigma^2 n\right] = \frac{\sigma^2}{n}$$

Using Chebyshev Inequality:

$c = \varepsilon$

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

Take $X = M_n$, for any $\varepsilon > 0$

$$P(|M_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

as $n \to \infty$

$$\boxed{P(|M - \mu| > \varepsilon) \longrightarrow 0}$$

Imp

The Central limit theorm (CLT):

Let $X_1, X_2 \ldots X_n$ be a sequence of independent identically distributed random variables with common mean μ & varience σ^2

$$\boxed{Z_n = \frac{X_1 + X_2 + X_3 \cdots + X_n - n\mu}{\sigma\sqrt{n}}}$$

Then CDF of Z_n converges to standard normal CDF

Explanation: $E[Z_n] = \frac{1}{\sigma\sqrt{n}}\left[\underbrace{E[X_1 + \cdots + X_n]}_{n\mu} - n\mu\right]$

$\Rightarrow E[Z_n] = 0$

$$Var[Z_n] = \frac{1}{\sigma^2 n}\left[Var(X_1) + Var(X_2) \cdots + Var(X_n)\right]$$

$$= \frac{\sigma^2 n}{\sigma^2 n} = 1.$$