11] Random Variables

*) A random variable is a variable whose Value is determined by the outcome of a random experiment.

*) It assigns a number to each ~~out~~ outcome in the sample space.

Ex:- i) Roll a fair dice.

- Sample space = {1, 2, 3, 4, 5, 6}
- Random variable "X" = the number showing on the die.

$$X \in \{1, 2, 3, 4, 5, 6\}$$

2) Toss two coins

- Sample space = {HH, HT, TH, TT}
- Random variable "Y" = the number of heads

$$Y \in \{0, 1, 2\}$$

Types of random variables

1) Discrete RV = Values are countable (ex: above)

2) Continuos RV = values are uncountable, usually measured (ex: time, height)

Date ___ / ___ / ___ №

12] Probability Distributions (Discrete)

*) A discrete probability distribution gives the probabilities of all possible values of a discrete random variable.

Ex :- Rolling a die → outcomes $\{1, 2, 3, 4, 5, 6\}$, each with probability $1/6$

Probability Mass Function (PMF)

The PMF is the function that assigns a probability to each outcome of a discrete random variable.

$$P(X = x) = f(x) \qquad \Sigma f(x) = 1$$

ex :- For a fair die, $P(X = 3) = 1/6$

*) $0 \leq P(X = x) \leq 1$

*) $\Sigma_x \, P(X = x) = 1$ (all probabilities add up to 1)

13] Binomial Distribution

A binomial distribution is a probability distribution that models the number of succeses in a fixed number of independent trials, each with the same probability of success.

• If flip a coin 10 times, what's the probability of getting exactly 6 heads?

Conditions for Binomial Distribution

1) The experiment consists of n independent trials
2) Each trial has two outcomes, Success (p) or Failure ($q = 1-p$)
3) We are interested in the number of successes (k) in trials.

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

X = random variable (number of success)
n = number of trials
k = number of successes
p = probability of success in a single trial
$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ ← number of ways to chose k success

Ex:- 1) Flip a coin 3 times ($n=3$, $p=0.5$) what's the probability of getting exactly 2 heads ($k=2$)

$$P(x=2) = \binom{3}{2} (0.5)^2 (0.5)^1$$

$$= 0.375$$

Date / /   №

Ex 2] A student guess on 5 questions, each with 25% of being correct ($n=5$, $p=0.25$)
Exactly 2 correct

$$P(X=2) = \binom{5}{2}(0.25)^2 0.75^2$$

$$= 0.2637$$

14] Bernouli Distribution

*] A bernouli distribution models a random experiment with only two outcomes:

- Success (1) with probability (p)
- Faliure (0) with probability ($1-p$)

PMF =>

$$\boxed{P(X=x) = \begin{cases} p & \text{if } x=1 \\ 1-p & \text{if } x=0 \end{cases}}$$

*] Binomial distrubition = Sum of multiple Bernouli trials (n trials)

15] Continuos probability distribution

*] Unlike discrete Variables, we can't assign a probability to a single exact value because there are infinitely many possible Values

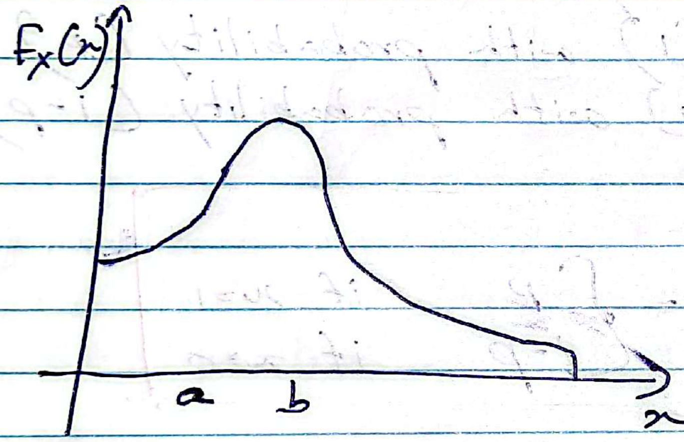
*] Ex:- Height of adults in cm could be any number from 140 to 200

16] Probability density function ($f_X(x)$)

*] Tell you the rate you accumulate probability around each point.

*] Only defined for continuos variables

*] $f_X(x)$ itself is not a probability, but the area under the curve between two values give the probability



$P(a<X<b)$ = area under $f_X(x)$

*] Defined for all numbers

*] $f_X(x) \geq 0$

*] Area under $f_X(x) = 1$

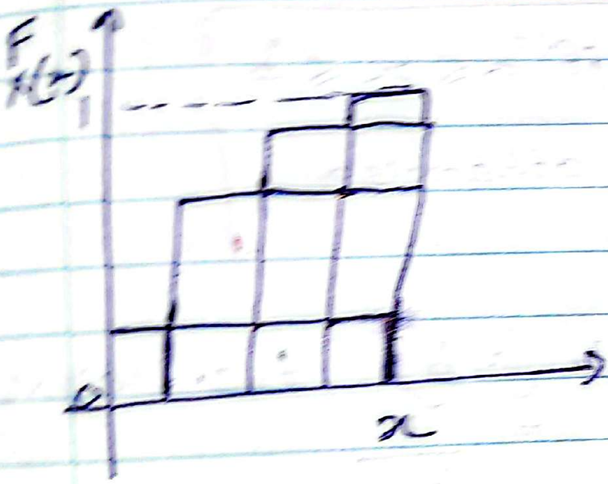
PMF : $P_X(x) = P(X=x)$

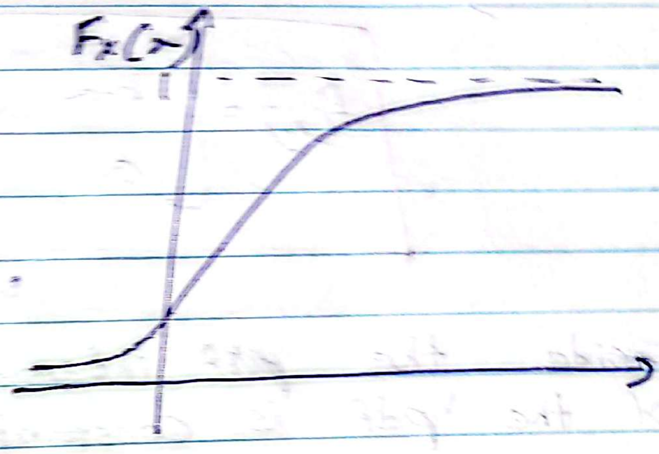
PDF : $f_X(x) = P(X=x) = 0$

*] Cumulative Distribution Function ($F_x(x)$)

*] Tells the probability that a continuos random variable is less than or equal to a value

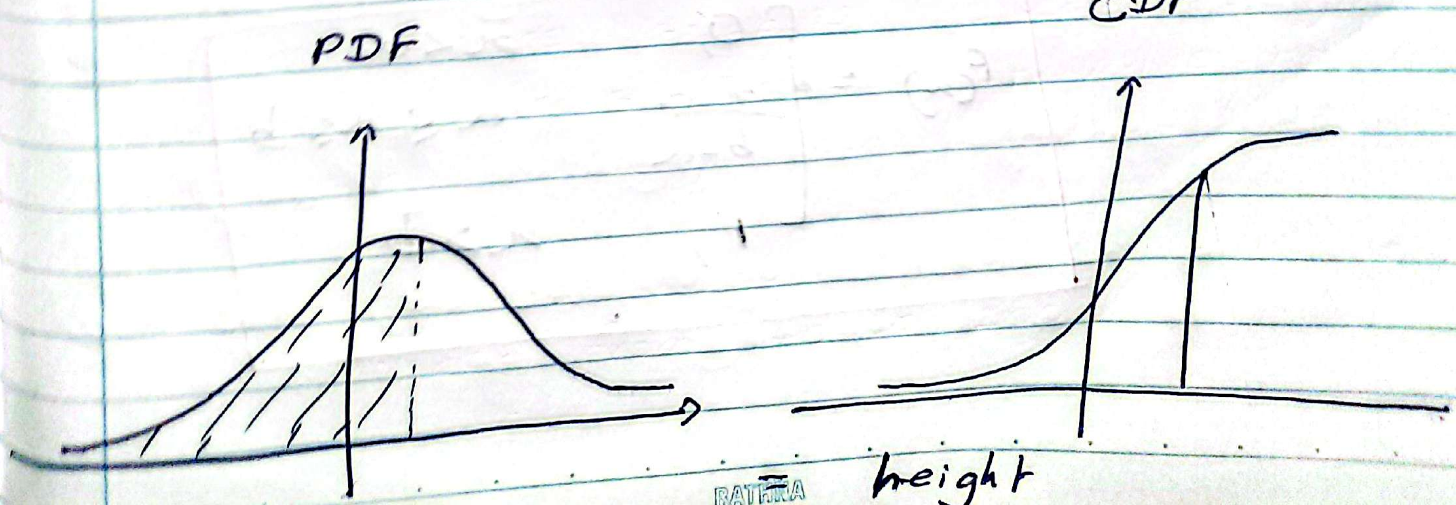
*] Shows how much probability the variable has accumulated until a certain value.





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

*] $0 \leq F_X(x) \leq 1$
*] left endpoint is 0
*] Right endpoint is 1
*] Never decreases



18] Uniform distribution (Continuos)

*) An uniform distribution is the simplest type of continuos distribution

- •) Every value in the interval (a, b) is equally likely

PDF =)

$$\boxed{f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}}$$

*) Inside the ~~pdf~~ interval (a, b) the "height" of the pdf is constant $= \frac{1}{b-a}$

*) Outside the interval, probability is 0

*) This ensures that the total area under the pdf = 1

CDF =)

$$\boxed{F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1, & x > b \end{cases}}$$

Date ___/___/___

№ ____

*) a is the minimum $(x < a) \Rightarrow p = 0$
*) b is the maximum $(x > b) \Rightarrow p = 1$

*) For any between $(a \leq x \leq b)$

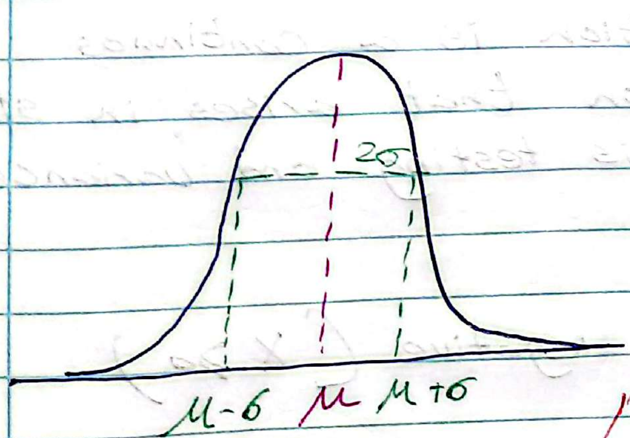
$$\boxed{p = \frac{x - a}{b - a}}$$

•) PDF ⇒ At any point in $[a, b]$, the likelihood density is the same

•) CDF ⇒ As you move on from a to b, probability accumulates steadily from a to b

19) Normal Distribution

*) The normal distribution is the most common distribution in real life it's bell-shaped and symmetric around the mean



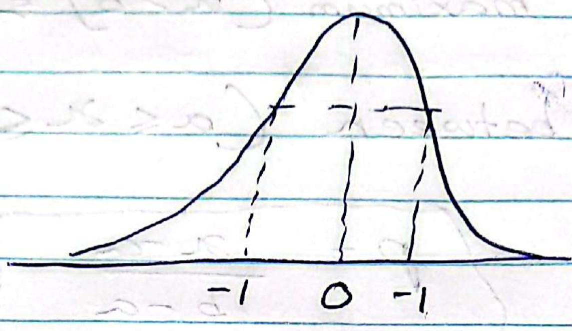
σ = spread of the bell (standard deviation)

μ = Center of the bell (Mean)

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

Scaling Constant

Standard Normal Distribution



$\mu = 0$

$\sigma = 1$

$X \sim \mathcal{N}(0, 1^2)$

*) To standardization any distribution

$$\boxed{Z = \frac{X - \mu}{\sigma}}$$

*) Z has a standard normal distribution

*) Standardization is crucial to compare variable of different magnitudes

20] Chi Squared Distributions

Chi squared distribution is a continues probability distribution that arises in stats, especially in hypothesis testing and variance analysis

*) It is always non negative $(X \geq 0)$

*) It is asymmetric (right skewed), but becomes more symmetric as degrees of freedom increase

Date ___ | ___ | ___ №___

$$\boxed{X \sim \chi^2(k) \quad \text{where } k \text{ is the degrees of freedom (df)}}$$

•)- If you have k independent standard normal random variables $Z_1, Z_2, \ldots Z_k \sim N(0,1)$ then the sum of their squares

$$X = Z_1^2 + Z_2^2 + Z_3^2 \ldots Z_k^2$$

follows a chi-squared distribution with k degrees of freedom.

*) Intution: It measures "squared deviations" from expected values - that's why it's used in goodness of fit tests.