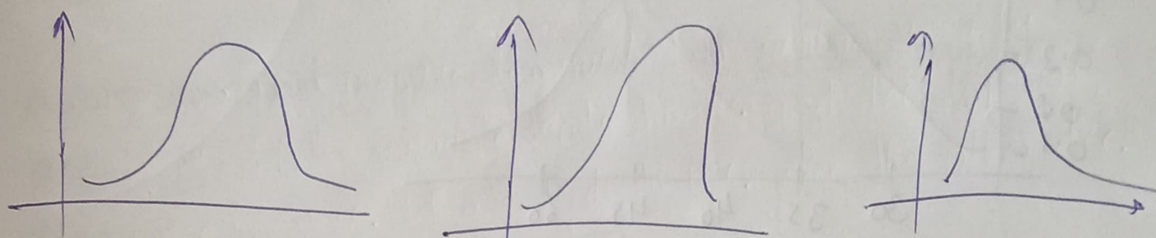
Probability Distribution function

→ PDF describes how the probabilities are distributed over the values of a random variable.

⇒ Age = { . ____ } → Continuous random variable



① Probability Mass function (PMF) → Discrete random values

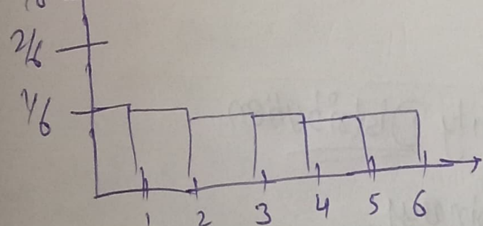
② Probability Density function (PDF) → Continuous random variables

---

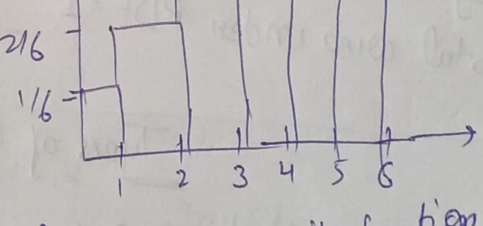
PMF

* Rolling a dice ⇒ $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$

PMF



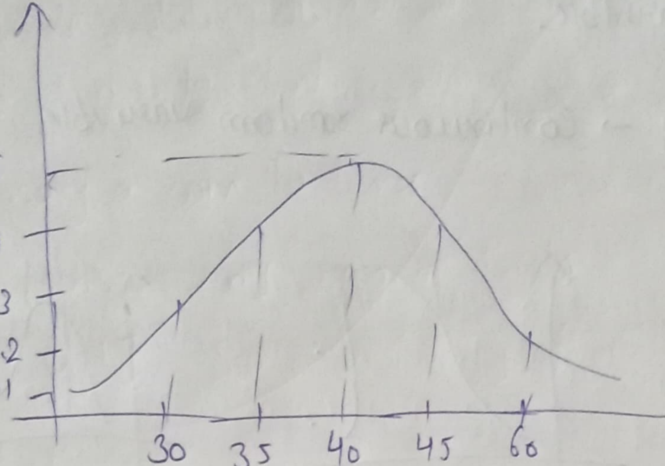
CDF



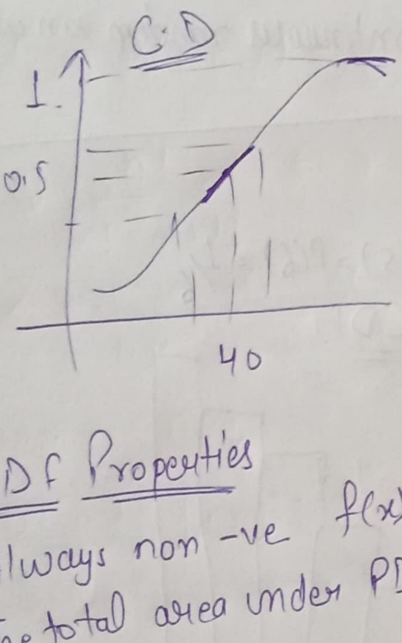
Commulative Density function

PDF

PD



PD is Gradient of Cummulative Density function



PDF Properties

→ Always non -ve $f(x) \geq 0$ ex.

→ The total area under PDF = 1. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Types of Probability Distribution

1. Bernoulli Distribution → Outcomes are binary
   pmf = Discrete random variable
2. Binomial Distribution :→ pmf
3. Normal / Gaussian Distribution → PDF
4. Poisson Distribution = Pmf
5. Log Normal Distribution = PDF
6. Uniform Distribution = Pmf

Bernoulli ① Distribution

⇒ The Simplest discrete probability distribution. It represents the probability distribution of a random variable that has exactly two possible (0/1).

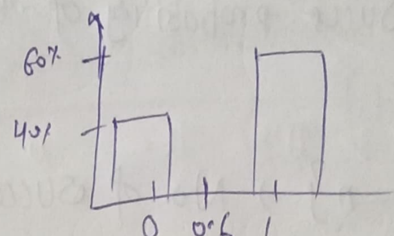
Parameters

→ $0 \le p \le 1$, $q = 1-p$, $k = \{0,1\}$ = 2 outcomes.

⇒ A company launch a phone, probability it was liked or not: 1

1 $P(\text{like}) = 60\%$

0 $P(\text{not like}) = 40\%$



⇒ **PMF**: $p^k \times (1-p)^{1-k}$

$P(k=1) = (0.6)(1) = 0.6 = p$

$P(k=0) = (1)(0.4) = 0.4 = q = (1-p)$

⇒ **Mean of BD**

$$E(x) = \sum_{i=0}^{k} k \cdot p(k) \qquad k = \{0,1\}$$

$= (0)(0.4) + (1)(0.6)$

$= 0.6 = p.$

⇒ **Mode** of I

$p > q$ = p will be mode

else q will be mode

Median of BD

$$\text{Median} = \begin{cases} 0 & p < 1/2 \\ [0,1] & p = 1/2 \\ 0 & p > 1/2 \end{cases}$$

Variance

$\sigma^2 = pq$

SD

$\sigma = \sqrt{pq}$

Binomial Distribution

⇒ In probability theory and stats, the BD with parameter n and p, is the discrete probability distribution of the number of succers in a sequence of n independent experiments, each asking yes or no question

* Tossing a coin 10 Times.

$n \in [0,1,2, -- n] \Rightarrow$ no. of Trials

$p \in [0,1] \rightarrow$ Succs probability of each trial

$q = 1-p$

$k \in \{0,1,2, -- n\} \Rightarrow$ No of successes.

PMF :→ $P(k,n,p) = {}^{n}C_{k}\, p^{k} (1-p)^{n-k}$

$${}^{n}C_{k} = \frac{n!}{k!\,(n-k)!} \qquad \in k = 0,1,2, \cdot\cdot - n.$$

Mean $= np$

Variance $= npq$

S.D. $(\sigma) = \sqrt{npq}$

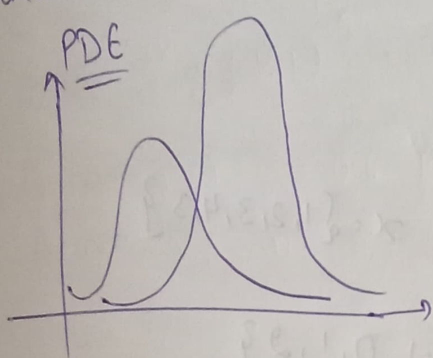
Poisson Distribution

→ It is a discrete probability distribution that express the probability of a given number of events ~~~~ occurring in a fixed interval of time, if these events occur with a known constant mean rate and independently of the time since the last event.

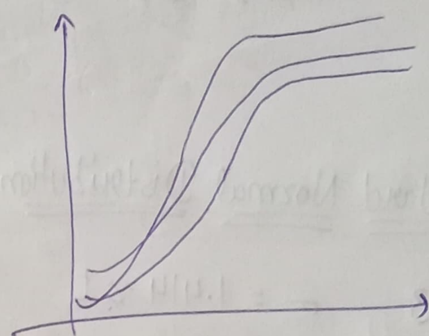
Normal / Gaussian Distribution

⇒ It is a type of continuous probability distribution for a ~~~~ real-valued random variable.

PDE



Cumulative distribution fun^n



μ = median = mode.

Notation $N(\mu, \sigma^2)$

$\mu \in R$

$\sigma^2 \in R \geq 0$ = variable

$x \in R$

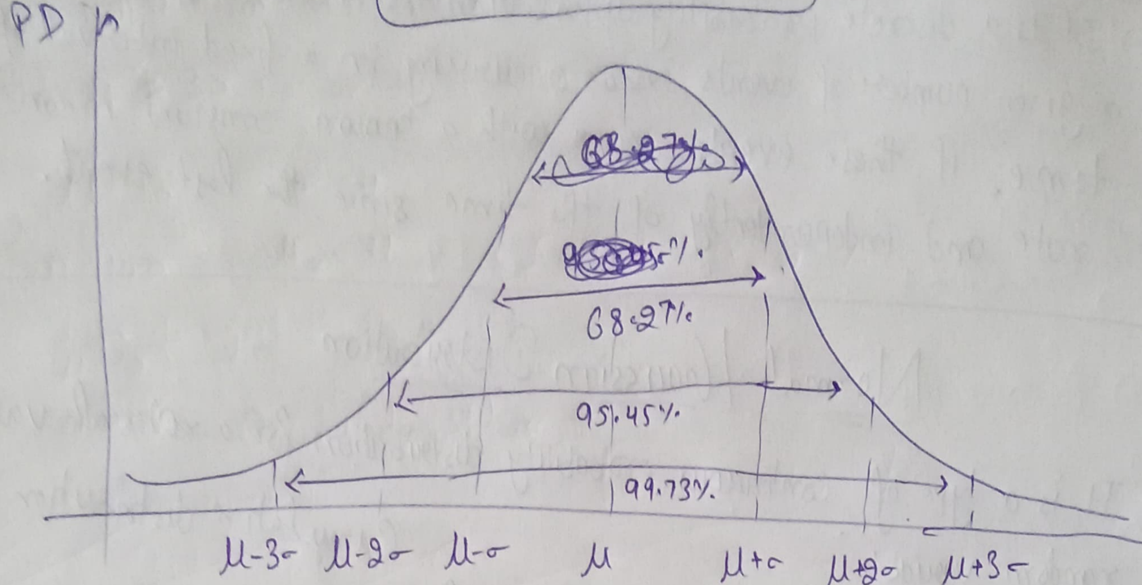
$$\boxed{PDF = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x_i - \mu}{\sigma}\right)^2}}$$

$$\boxed{\mu = \sum_{i=1}^{N} \frac{x_i}{N}}$$

$$\boxed{\sigma^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{N}}$$

$$\boxed{\sigma = \sqrt{\sigma^2}}$$

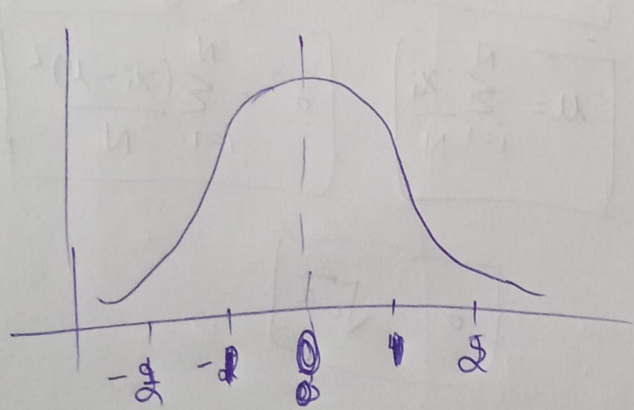
68 -95-99.7 Rule



1. Standard Normal Distribution

$\mu = 3$ $\quad \sigma = 1.414 \approx 1$ $\quad N=5$ $\quad x = \{1,2,3,4,5\}$

$Z\text{-score} = \dfrac{x_i - \mu}{\sigma} \Rightarrow y = \{-2,-1,0,1,2\}$



Confidence Interval (CI) $= \mu_{\bar{x}} \pm Z \cdot \sigma_{\bar{x}} = \mu_{\bar{x}} \pm Z \dfrac{\sigma}{\sqrt{n}}$