[Distribution of Data

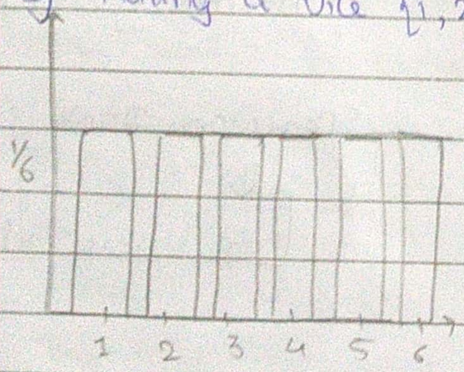
# Probability Distribution function

| Probability density function (pdf) | Probability mass function |
|---|---|
| - continuous value | - Discrete value |
| - eg: Age | - Eg: No. of Bank Account |

- Whenever we draw probability density function, initially there will be histogram & when we smoothen that histogram we get probability density function for continuous graph.

→ Probability Mass function [pmf]

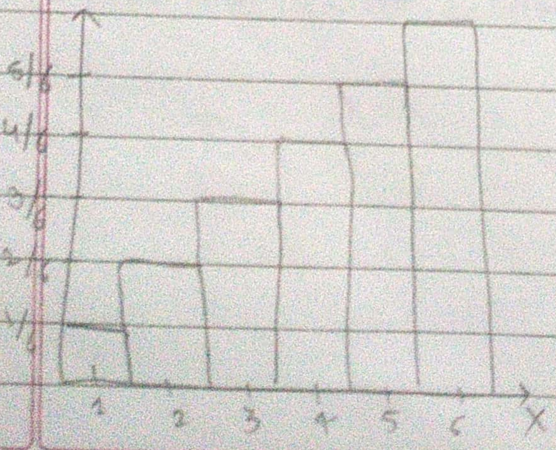
Eg: Rolling a Dice {1, 2, 3, 4, 5, 6}




$$Pr(x \le 4) = Pr(x=1) + Pr(x=2) + Pr(x=3) + Pr(x=4)$$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

→ Cummulative Distribution function [cdf]



- Probability density of pdf ⇒ Gradient or derivative of cdf

  → pdf is derivative of cdf
  → cdf is integration of pdf.

→ Types of Probability Distribution.

1) Bernoulli Distribution.
- Discrete random variable {pmf}
- outcomes are binary
- Eg. Tossing a coin {H, T}

  $Pr(H) = 0.5 = p$ , $Pr(T) = 0.5 = 1-p = q$

- $0 \le p \le 1$ , $q = 1-p$

  $k = \{0, 1\} \Rightarrow$ outcomes.

- pmf $\begin{cases} q = 1-p & \text{if } k=0 \\ p & \text{if } k=1 \end{cases}$

- $pmf = p^k * (1-p)^{1-k}$

- Mean of bernoulli distribution

$$E(k) = \sum_{i=0}^{k} k \cdot p(k)$$

$Pr(k=1) = 0.6 \to p$
$Pr(k=0) = 0.4 \to 1-p$

$$= [0 * 0.4 + 1 * 0.6]$$
$$= 0.6$$

$\therefore E(k) = p$

2) Binomial Distribution

- Discrete random variable
- every experiment outcome is binary
- These experiment is performed for n trials

Eg. Tossing a coin 10 times

- Notation : $B(n, p)$

$Pr(H) = 0.5 = p$

$Pr(T) = 0.5 = q$

- Parameters : $n \in \{0, 1, 2, 3, \ldots\} \Rightarrow$ no. of trials

$p \in \{0, 1\} \rightarrow$ success probability

$q \in 1 - p$

support : $k \in \{0, 1, 2, \ldots n\} \rightarrow$ number of successes

- pmf

$$Pr(k, n, p) = {}^nC_k \, p^k (1-p)^{n-k}$$

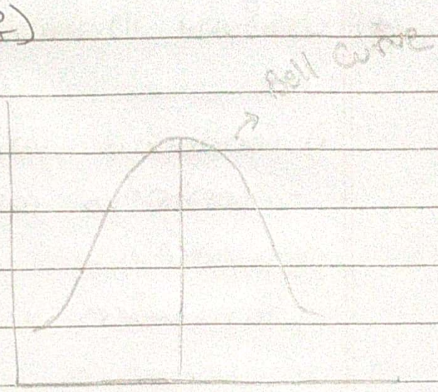
3) Poisson Distribution

- Discrete random variable (pmf)
- Describe the number of event occuring in a fixed time interval.

most important

**4) Normal / Gaussian Distribution (pdf)**

- continuous random variable
- Graph is always symmetric

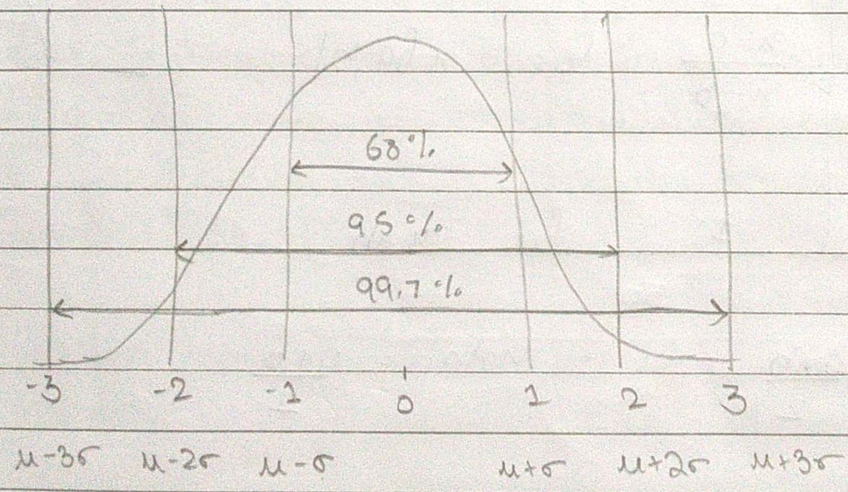



- Notation : $N(\mu, \sigma^2)$
- Parameters : $\mu \in R$
  - $\sigma^2 \in R > 0$ = variance
  - $x \in R$

- Pdf $= \dfrac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$

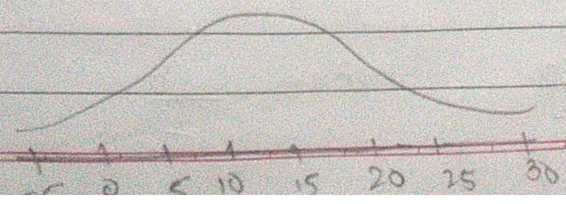
$\because \mu$ = Mean
$\sigma^2$ = variance
$\sqrt{\sigma^2}$ = std

→ Emperical rule of normal Distribution




$\mu - \sigma \le x \le \mu + \sigma \approx 68\%$ $1^{st}$ standard deviation
$\mu - 2\sigma \le x \le \mu + 2\sigma \approx 95\%$
$\mu - 3\sigma \le x \le \mu + 3\sigma \approx 99.7\%$

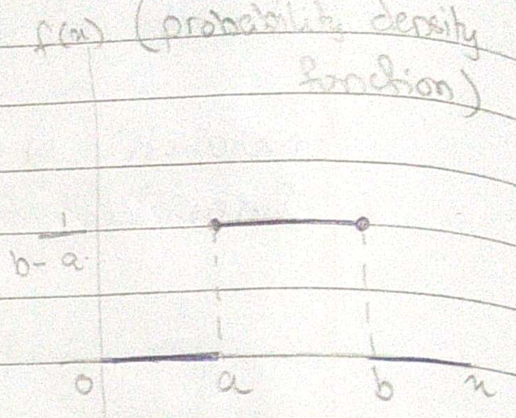
→ $\mu = 10$ $\sigma = 5$ ⇒ Normal Distribution



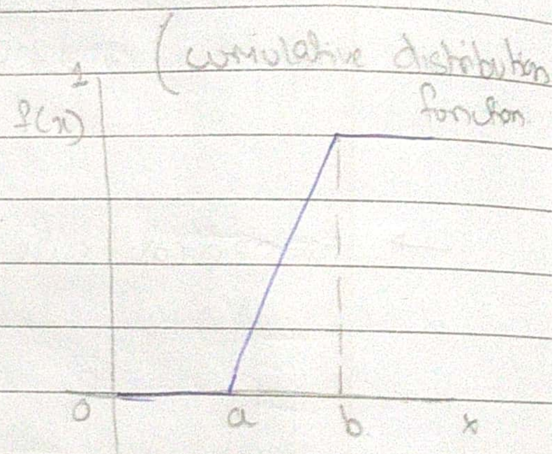
## ⊛ Uniform Distribution

1) Continuous Uniform Distribution [pdf]

- continuous random variable
- Notation = $U(a,b)$
- Parameters = $-\infty < a < b < \infty$




- pdf = $\begin{cases} \frac{1}{b-a} & \text{for } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$




- cdf = $\begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{x-b} & \text{for } x[a,b] \\ 1 & \text{for } x > b \end{cases}$

- Mean = $\frac{a+b}{2}$
- Median = $\frac{a+b}{2}$

2) Discrete Uniform Distribution [pmf]

- Notation $U(a,b)$
- Parameter $a, b$ $\boxed{b \geq a}$
- pmf = $\frac{1}{n}$

- Mean = $\frac{a+b}{2}$
- Median = $\frac{a+b}{2}$

* Standard Normal Distribution & Z-score

Z - Stats

- Whenever there is Standard Normal Distribution, we have Mean ($\mu$) = 0 & standard deviation ($\sigma$) = 1.

To convert

we use

- Normal Distribution → $\boxed{\text{Z-score} = \frac{x_i - \mu}{\sigma}}$ → Standard Normal Distribution

↓

Transformation technique.

→ Normalization & Standarization.

**Normalization:**

It's also feature scaling technique but here we provide some range value whether it has to get converted to (0, 1), (-1, 1) or (0, n).

Eg: Used in Deep Learning technique

- min-max, scaler

**Standarization:**

If we want to scale down the data almost to the same unit using Z-score.

- It is feature scaling technique
- It is used ~~to~~ for training the model
- Eg: $\frac{x_i - \mu}{\sigma}$, $\mu = 0$, $\sigma = 1$

- feature transformation

| feature scaling | → | Model | $\xrightarrow{o/p}$ | Normal Data |
|---|---|---|---|---|

Same unit

Rescale unit