Probability Distribution 2

Agenda:

1) Population vs Sample
2) Sample Statistics
3) Point Estimates
4) Standard Error
5) Sampling Techniques
6) Uniform Distribution

* Population vs Sample.

1. Practicality
   - time Consuming
   - Cost in-effective
   - lots of resources

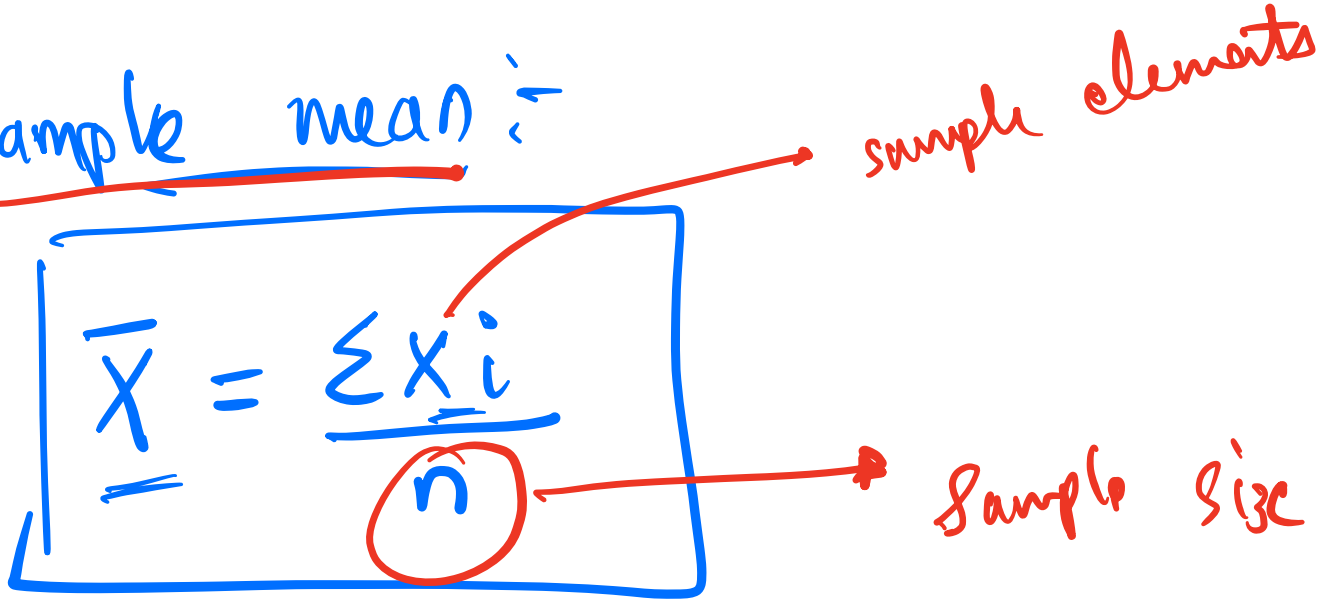
Sample Statistics :

1) Sample mean ($\bar{x}$)

2) " variance

3) " std. dev

Sample mean :



2) Sample Variance :- (σ^2)

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Bessel's correction

3) Sample Std. dev (σ)

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

(*) Sample Size (n) = 5

$$\text{Sample mean } (\bar{x}) = \frac{131+150+140+142+152}{5}$$

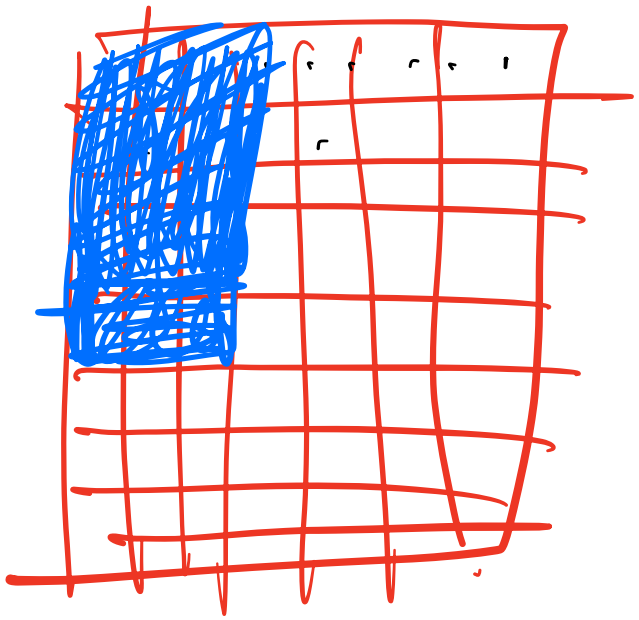
$$= 143 \text{ cm}$$

2) Sample variance (σ^2)

$$= \frac{(131-143)^2 + (150-143)^2 \cdots + (152-143)^2}{(5-1)}$$

$$= 71 \text{ cm}^2$$

Point Estimates :-



1) Sample mean ⟶ pop mean

2) Sample variance ⟶ pop variance

Sampling Techniques :-

1. Probability Sampling
2. Non-Probability Sampling

1. Probability Sampling :-

1. Simple Random Sampling.
2. Systamatic Sampling
3. Stratified Sampling
4. Clustered Sampling.

2. Non-Probabilistic

→ 1. Snowball Sampling.

* Simple Random Sampling :-

→ Population: 1 lakh

Sample = 1000 people.

$$= \frac{1000}{1 \text{ lakh}} \rightarrow 0.01$$

Two methods for Random Selection

① Lottery based.

Standard Error :

↳ Quantifies the variability across samples.

std-dev → variability within a sample

Std. error (SE) → variability across sample

$$SE = \frac{\sigma}{\sqrt{n}}$$

where SE → Standard error

σ → Population std. dev

n → Sample Size

$$SE = \frac{s}{\sqrt{n}}$$

s → Sample std. dev

Quiz 2:-

$SE = ?$

$n = 30$

$\bar{x} = 4$

$S = 0.13$

$$SE = \frac{S}{\sqrt{n}}$$

$$= \frac{0.13}{\sqrt{30}}$$

$$= 0.02$$

* Uniform Distribution (UD)

- Discrete
- Continuous

eg:- Discrete UD

Dice throw (6 faced)

2.3
2.7 2.35

$S = \{1, 2, 3, 4, 5, 6\}$

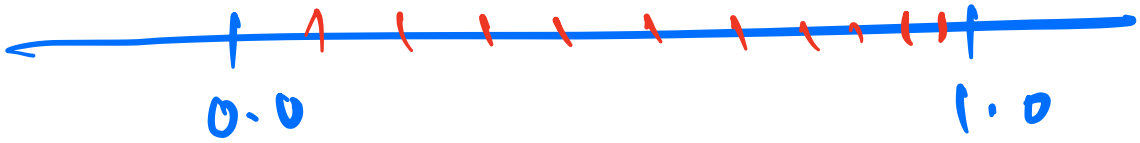
$$P(X) = \frac{1}{6} = \frac{1}{n}$$

n → no. of elems

$P \times n = 1$

$P = 1/n$

* Continuous UD



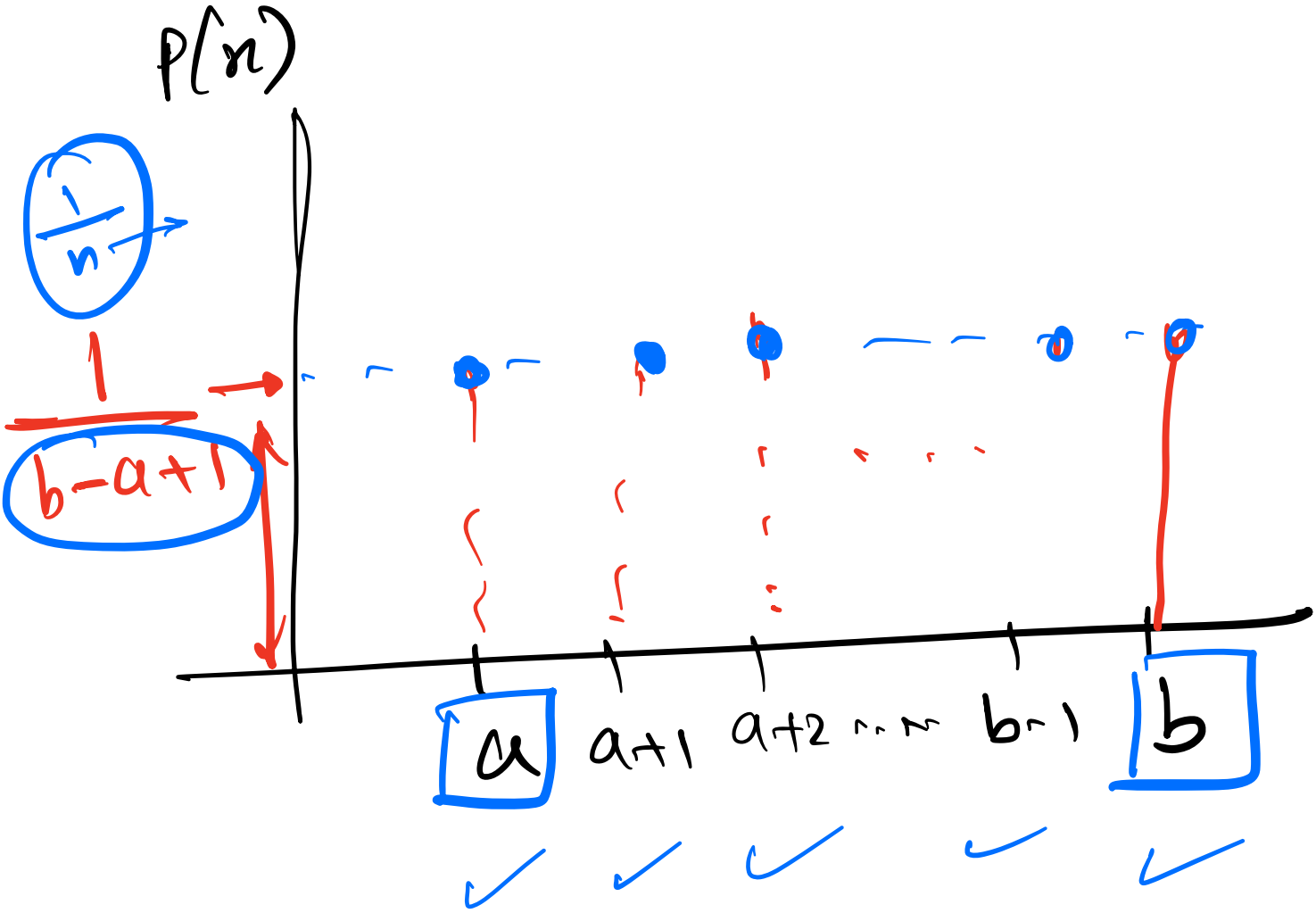
Discrete UD:

$$PMF(x) = PMF(X = x)$$

$$= \frac{1}{b - a + 1}$$

Prob Mass function

$[a, b]$ $\quad a \le x \le b$



$a \rightarrow 1$

$b \rightarrow 6$

$6-1+1$

$[1, 6] \rightarrow$

$1, 2, 3, 4, 5, 6$

$[a \quad b]$

$$b - a + 1$$

$[2, 8]$ → $8 - 2 + 1 = 7$

$2, 3, 4, 5, 6, 7, 8.$

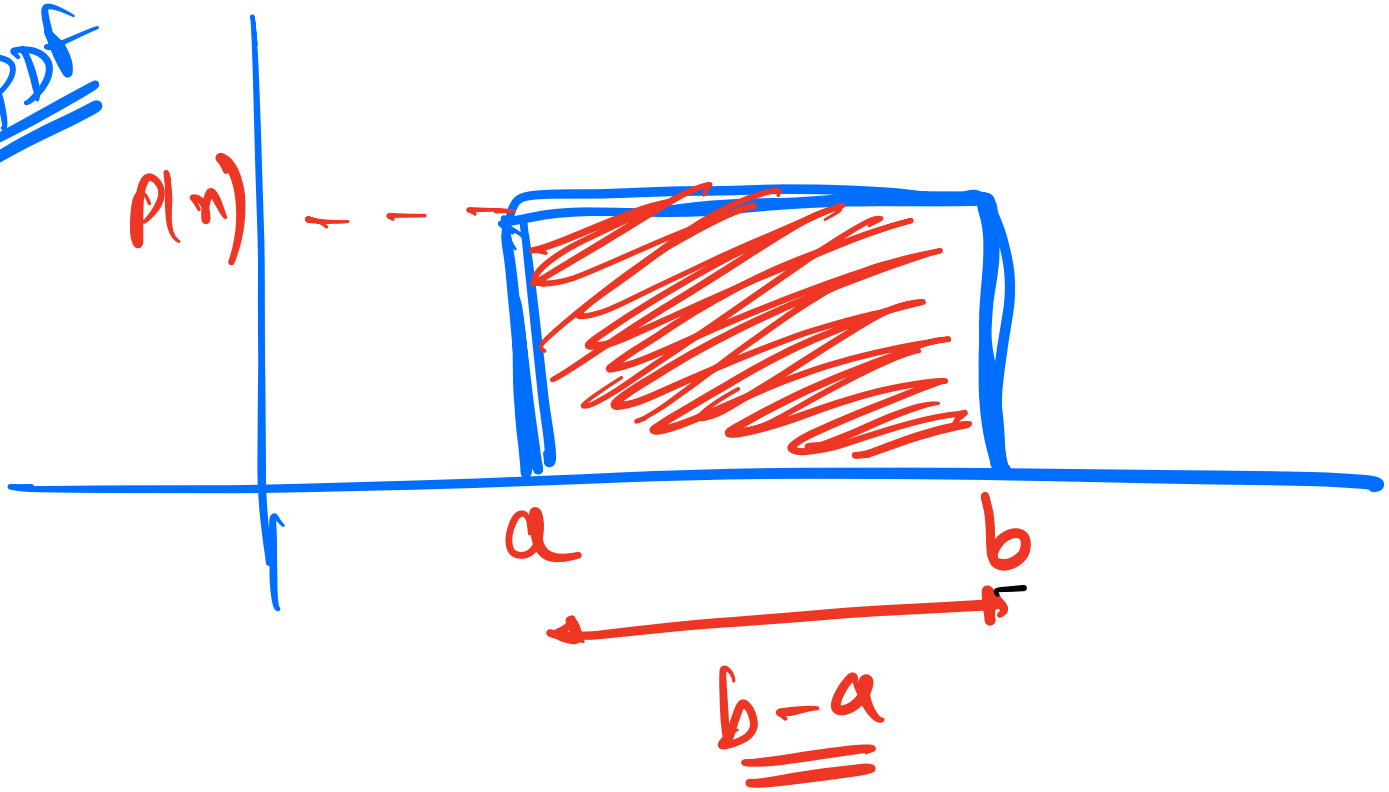
7

* Continuous UD

$$PDF(x) = \frac{1}{b-a} \quad \text{for} \quad \boxed{a \le x \le b}$$

$$= 0 \quad \text{(otherwise)}$$

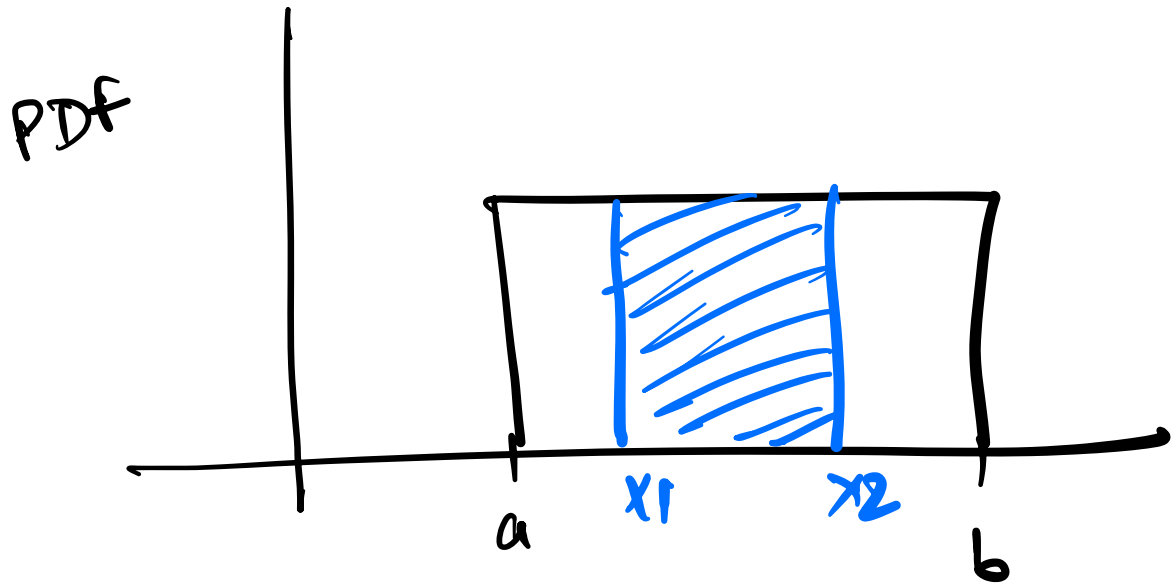
PDF



$$P(x) * (b-a) = 1$$

$$\boxed{P(x) = \frac{1}{b-a}}$$

(a) $P(x_1 < x < x_2)$



$$P(x_1 < x < x_2) = \frac{x_2 - x_1}{(b - a)}$$

Ques 3:-

$a = 100$

$b = 150$

$$P(120 < X < 130)$$

$$= \frac{130 - 120}{150 - 100}$$

$$= \frac{10}{50} = \boxed{0.2}$$

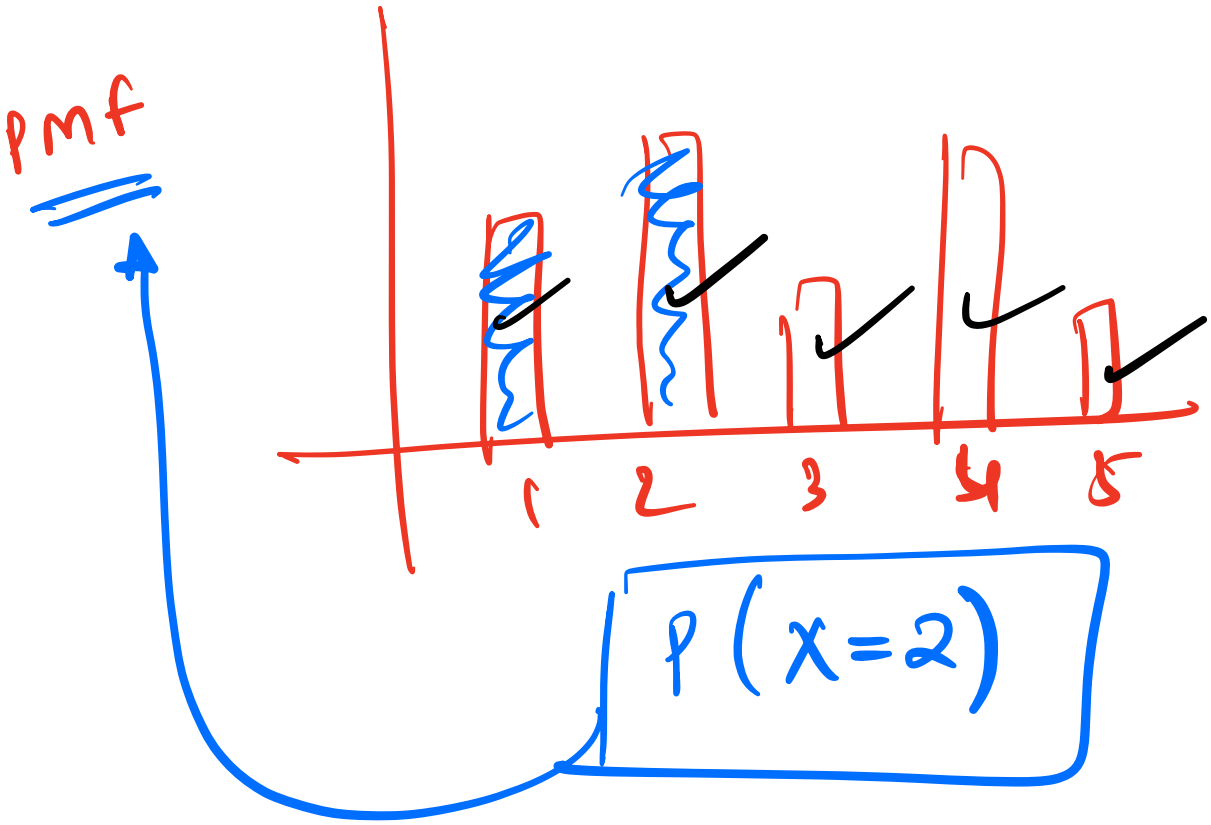
U D formulas

mean: $\frac{a+b}{2}$

median: $\frac{a+b}{2}$

variance: $\frac{(b-a)^2}{12}$

Std. dev: $\sqrt{\frac{(b-a)^2}{12}}$



P(x < 3)

cdf

or

pmf(X=1) + pmf(X=2)

(39) $P(X > 3) = 1 - P(X \leq 3)$