Descriptive Statistics.

1. Measures of Frequency
   - Counts, percentage, proportions
2. Measures of Central tendency
   - Mean, Median, Mode
3. Measures of dispersion
   - Variance, Standard deviation, range.
4. Measures of position.
   - Percentile, quartiles, deciles, ...

1 / 3

PowerPoint File Edit View Insert Format Arrange Tools Slide Show Window Help

Random Variable.

— Probability distributions.

Probability = Chance of an event occuring.

= 0 to 1 (0 – 100%).

= –ve probabilities not possible.

P(H) = $\frac{1}{2}$ = # No. of Fav. outcomes / Total No. of possible outcomes. → Sample space.

↳ Experiment

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

$P(X=H) = \frac{1}{2}$

Tossing a Coin = [H, T] → discrete Random Variable

Rolling dice = [1, 2, 3, 4, 5, 6] → sample space contains only finite no. of options.

6

Prob (1) ⇒ 1/6.

2 / 6

PowerPoint File Edit View Insert Format Arrange Tools Slide Show Window Help

Continuous Random Variable ←

→ [150, 151, 152, 153]

120 190

X → Random Variable used to hold the height of student

[150, 150.01, 150.001, 153.5 . . . .]

↳ ∞

$$P(X = 153.5) = \frac{1}{\#\text{ total no. of possible outcomes}} = \frac{1}{\infty} = 0$$

$(153 < X < 154)$ → Finite Value

3 / 6

$X \rightarrow$ Can take a set of possible values

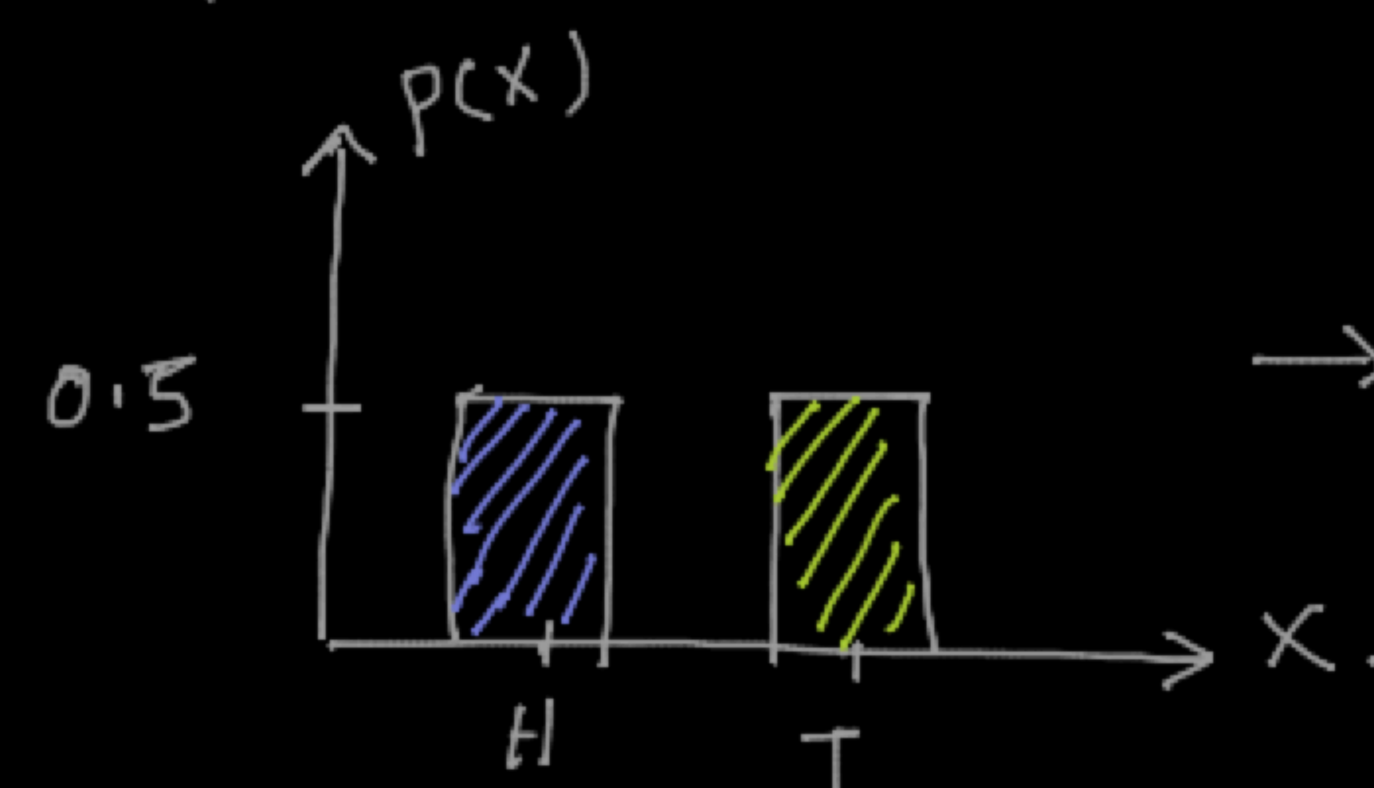
$\rightarrow$ Probability distribution of X.

Discrete Random Variable

Continuous random Variable.

$X = [H, T]$

$P(X=H) = 0.5$; $P(X=T) = 0.5$



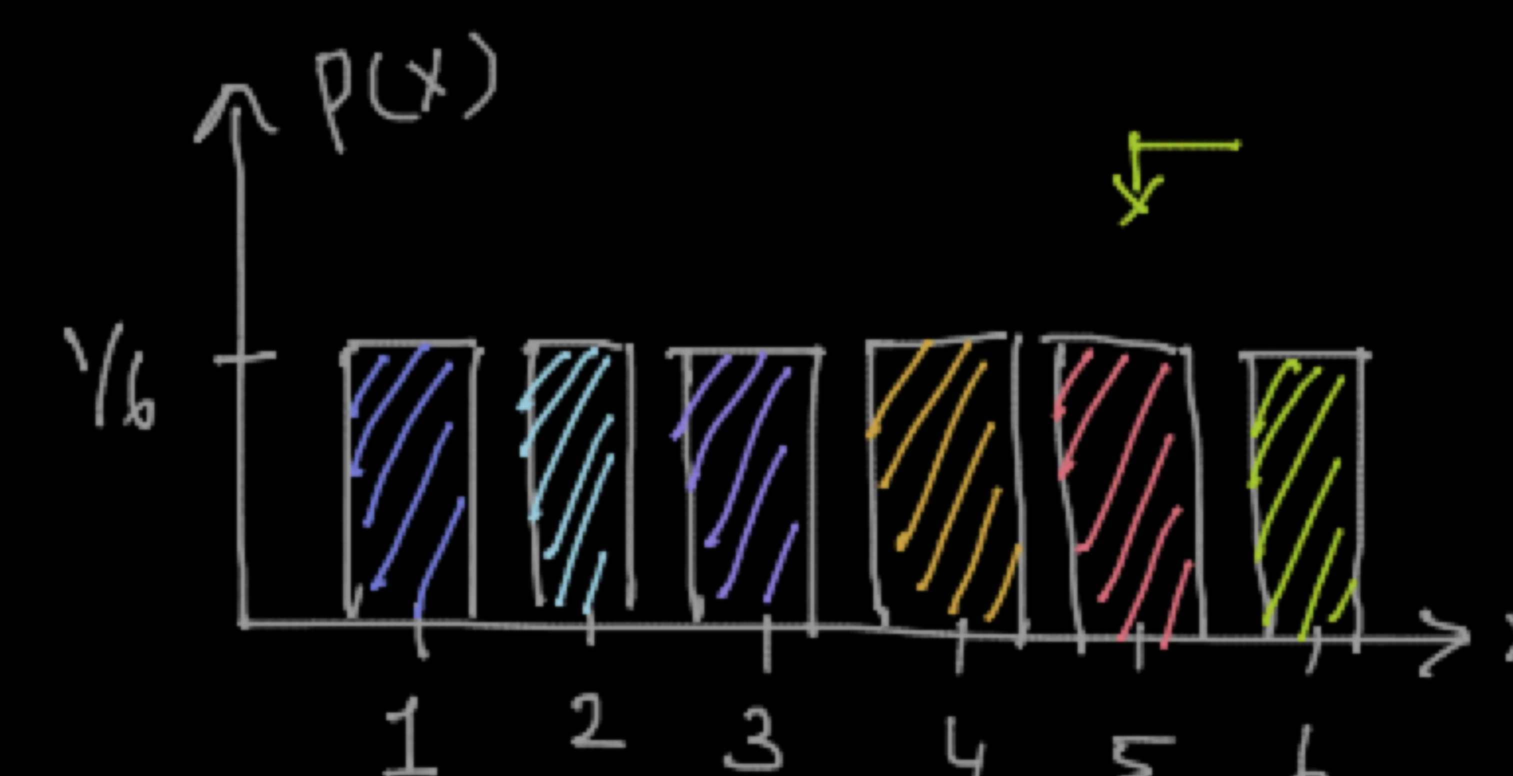
$\rightarrow$ uniform distribution $\leftarrow$

$\hookrightarrow$ Bernoulli distribution

$X = [1, 2, 3, 4, 5, 6]$

$P(X=1) = 1/6$; $P(X=3) = 1/6$

$P(X=2) = 1/6$; $P(X=4) = 1/6$



Uniform distribution

— Every possible outcome has equal prob. of occuring.

Bernoulli distribution

— repeat an expt. only ONCE

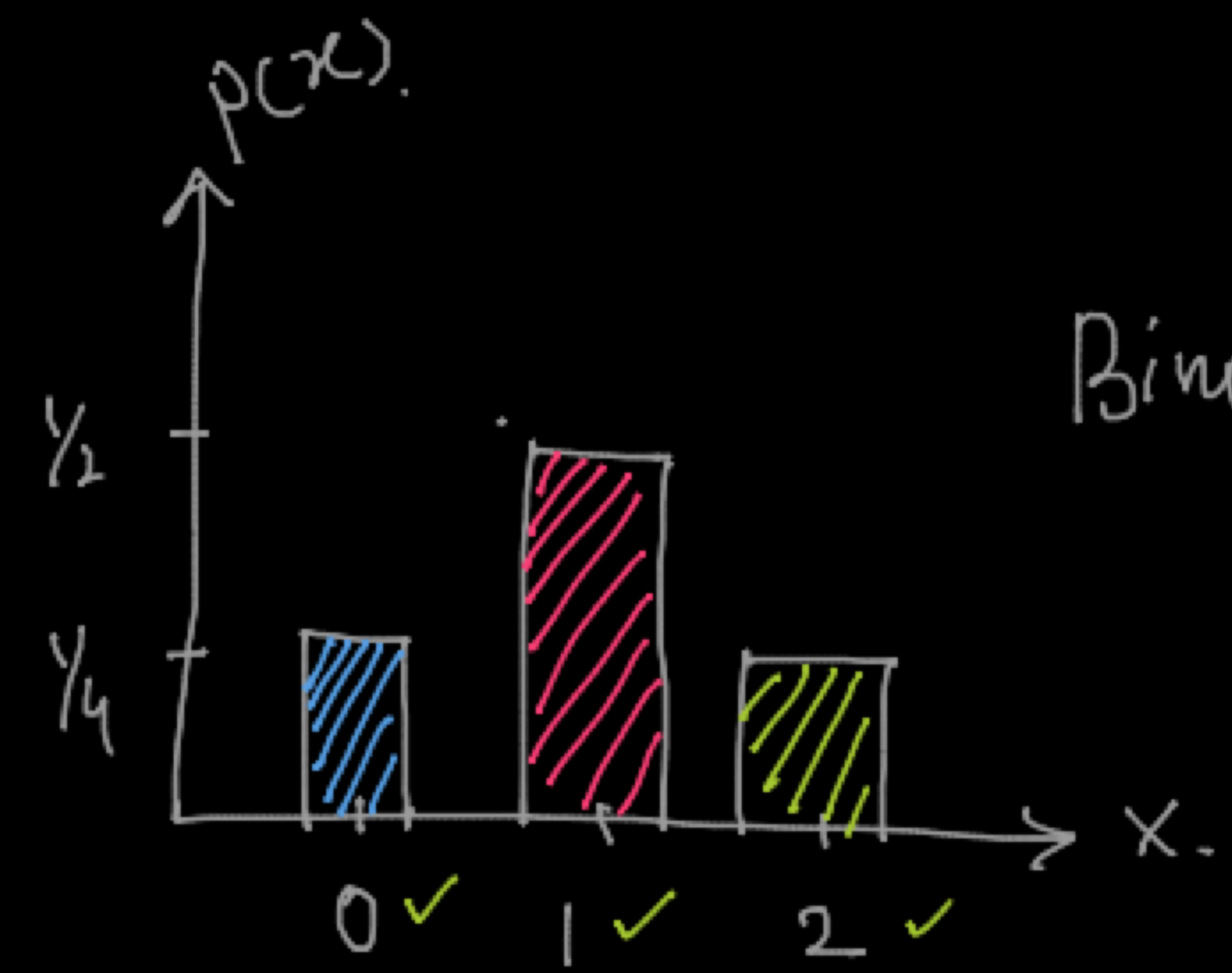
— There should be only 2 possible outcomes.

Binomial distribution.

— Toss a coin twice, thrice, . . .

4

H H → 1/4
H T
T H } 2/4
T T 1/4



$x=0;\ x=1,\ x=2,\ x=3$

Binomial distribution

Prob. dist. this X

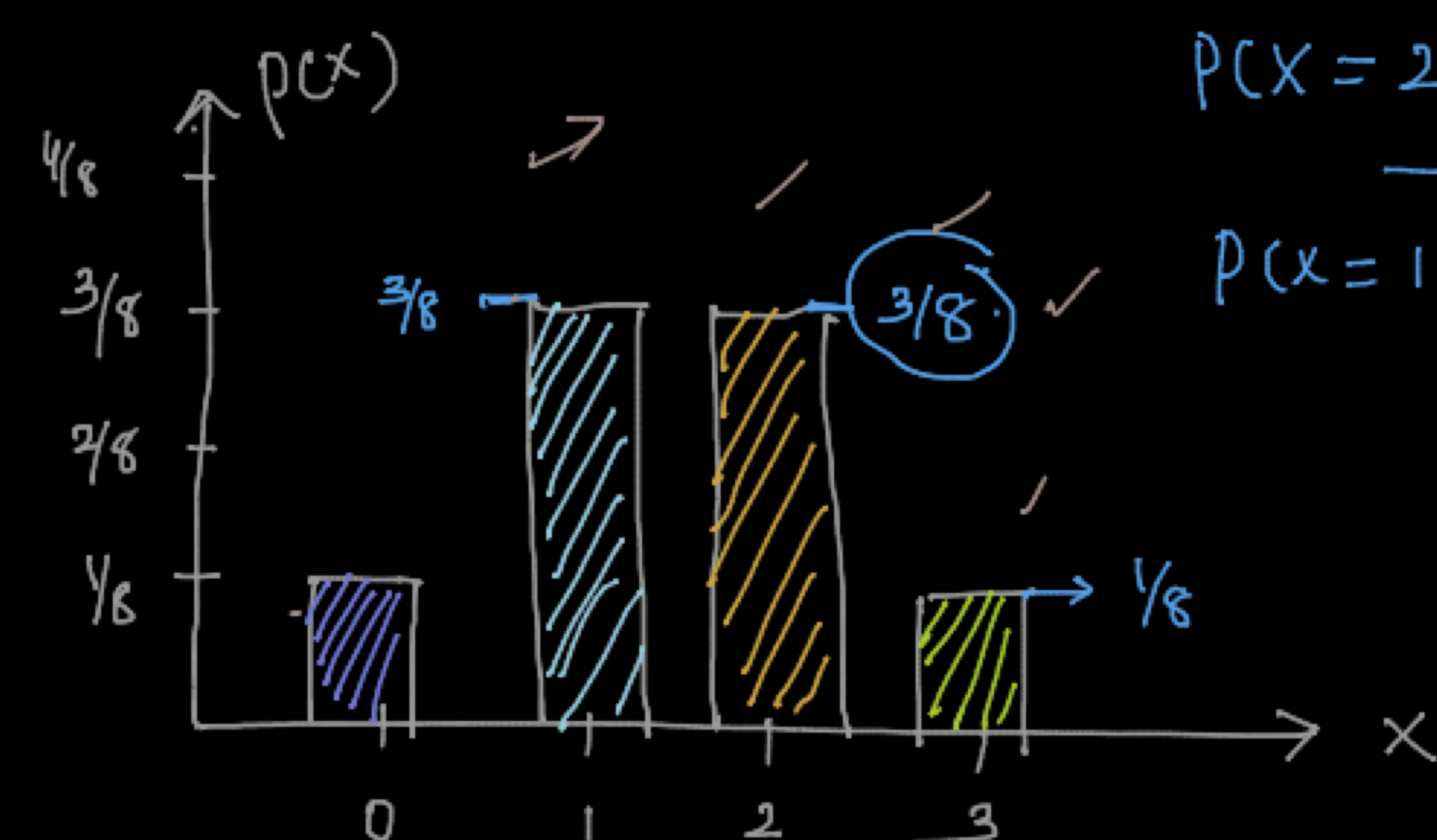
$\boxed{50\text{ times}, \ 33\text{ Heads ?}}$

$$= \frac{n!}{(n-x)!\,x!} \cdot p^{x} \cdot q^{(n-x)}$$

n – no. of times Expt.
x → No. of success.
p → prob. of success (0.5)
q → prob. of fail (0.5)

P(X = 0) P(X = 1) P(X = 2H) P(X = 3H)

H H H
H H T
H T H
H T T
T H H
T H T
T T H
T T T



P(X = 2H) → 3 times
P(X = 1H).

$n=3, \quad x=2$

$$P(x=2) = \frac{3!}{(3-2)!\,2!}(0.5)^{2}(0.5)^{(3-2)}$$

$$= \frac{3!}{1!\,2!} \times (0.5)^{2} \times 0.5$$

$$= \frac{3\times\cancel{2\times1}}{\cancel{2\times1}} \times (0.25)\times 0.5$$

$$= 3 \times \frac{1}{4} \times \frac{1}{2}$$

$$= 3/8$$

Random Variable

Probability distribution Fn.

Discrete

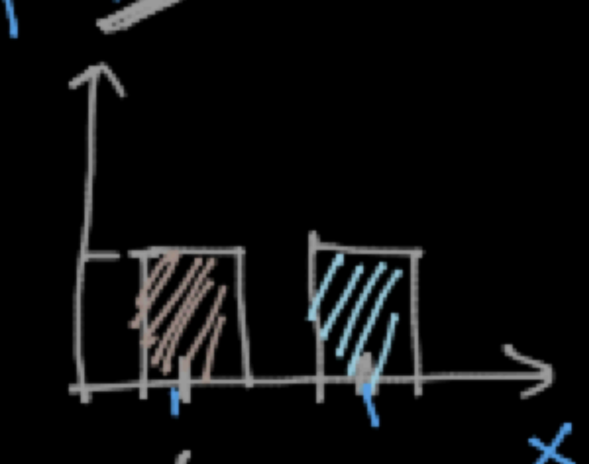
— only finite possible Values

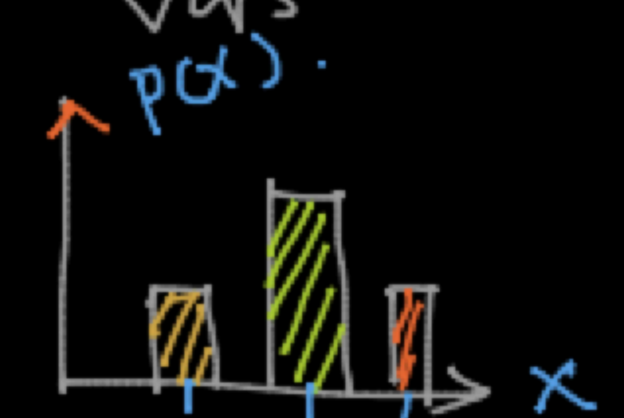
— prob. of x taking a specific value is finite.

$X = [H, T]$; $X = [1, 2, 3, 4, 5, 6]$

$P(X = H) = 1/2$ $P(X = 2) = 1/6$.

— Prob. dist. of discrete random vars





Bernoulli, Binomial, Uniform..

Probability Mass Fn (p.m.f)

$$P(x) = \frac{n!}{(n-x)!\,x!} \cdot p^{x} \cdot q^{(n-x)}$$

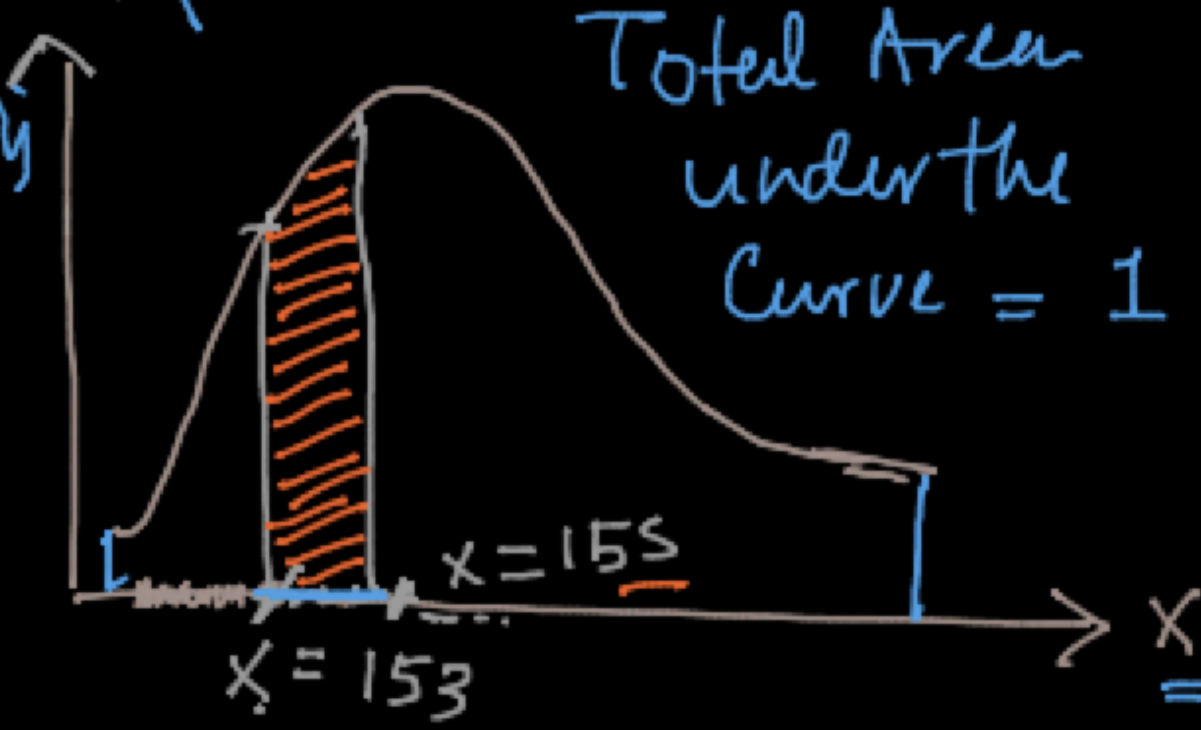
Continuous.

— infinite possible Values.

$P(X = 153) = 0$

— Probability density Fn. (p.d.f)

? p.d.f



Total Area under the Curve = 1

— Normal distr.

— SND (Z-dist)

— Student's (t-dist.)

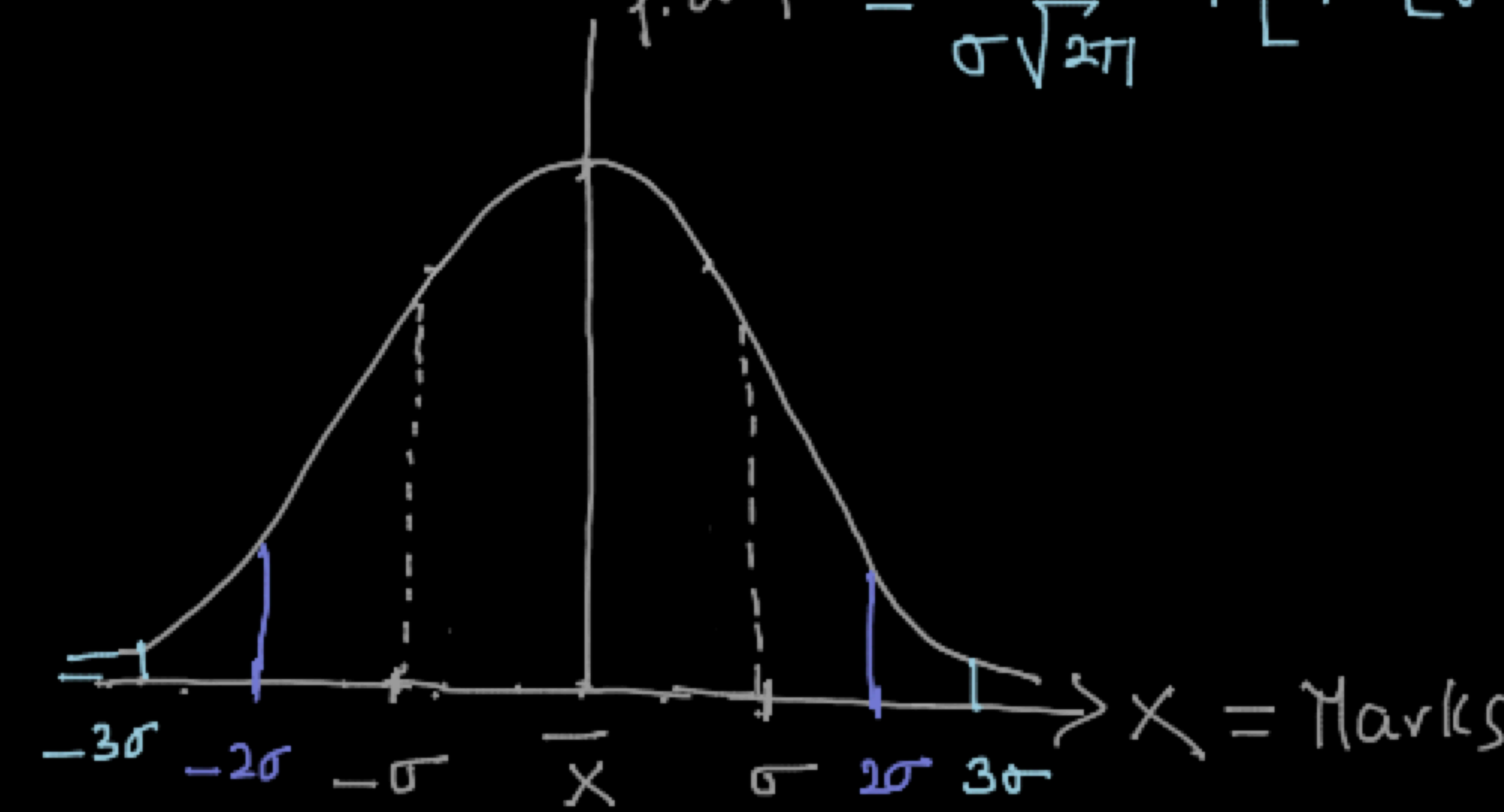
$P(X = 1)$

$$P(153 < X < 155) \Rightarrow \int_{x=153}^{155} p.d.f.$$

1. Normal Distribution

- Bell shaped
- Symmetric
- Mean/Median/Mode
- Empirical Rule.

$$\text{p.d.f} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right]$$



50.5 61 71.5

← 68% →

← 95% →

← 99.7% →

61 σ = 10.

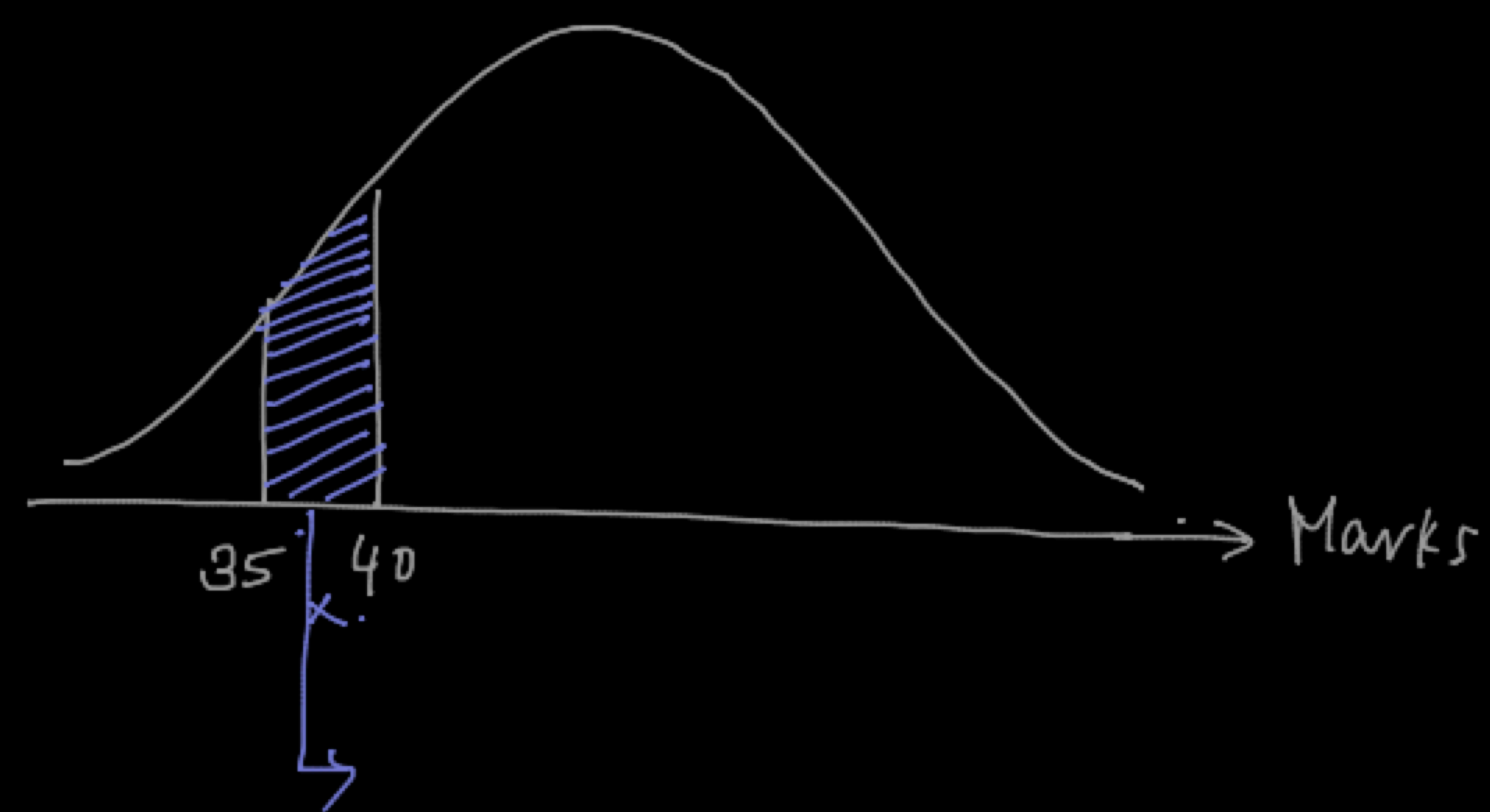
68% →	51	71
95% →	41	81
99.7% →	31	91

Marks (x)

→ 55
→ 43
→ 67
→ 73
51
48

$\bar{X} =$

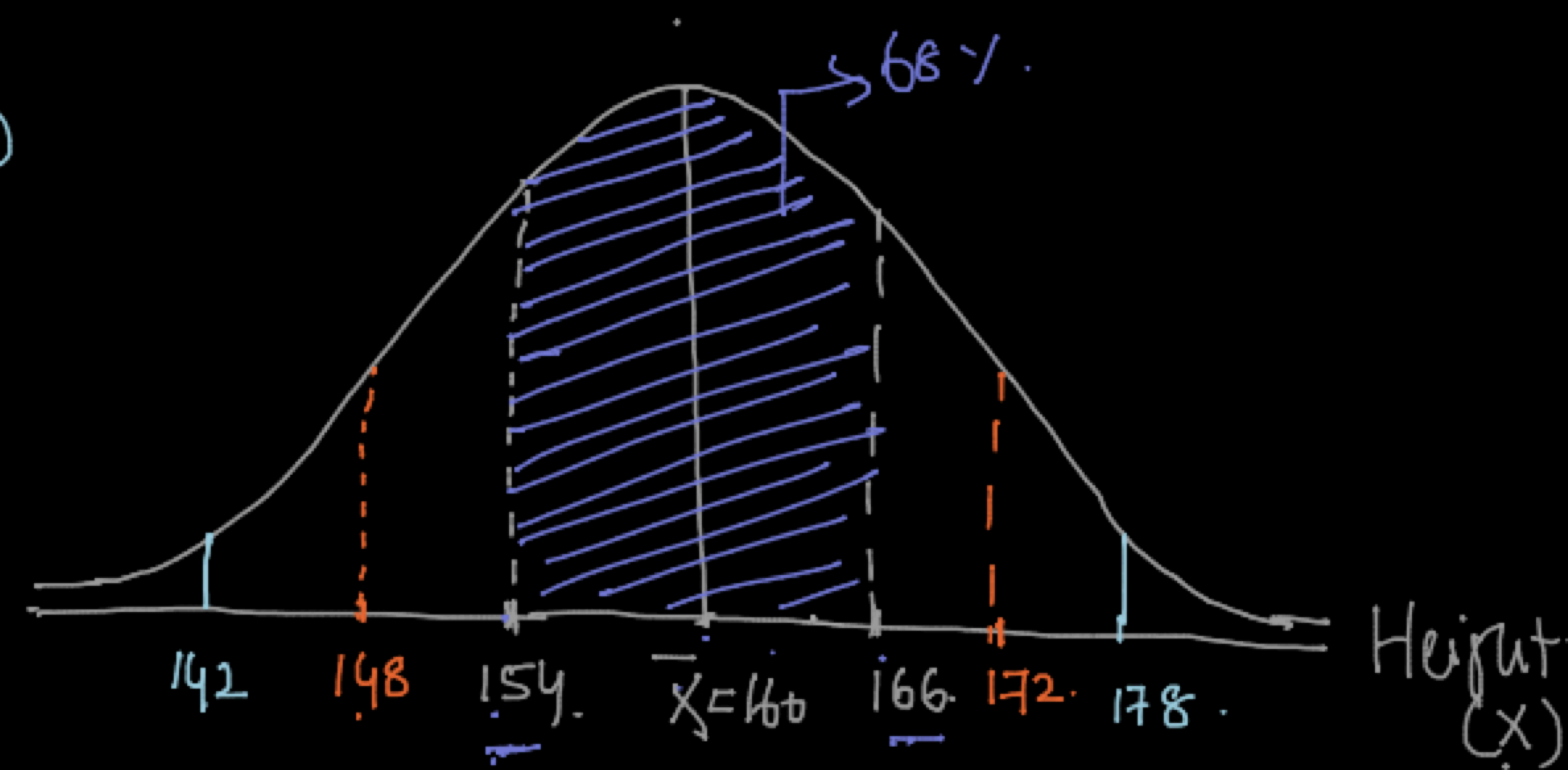
$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$



$P(35 < \text{Mark} < 40) =$

$\text{Height} \sim N(\mu, \sigma)$

$\sim N(160, 6)$



$\leftarrow P(148 < H < 172) \rightarrow$

95%

$\leftarrow$ 99.7% $\rightarrow$

Heights (X)

$\bar{x} = 160.$ $\sigma = 6$ cms.

$\boxed{n = 100} \rightarrow 68$ ✓

$\bar{x} \pm 2\sigma$

$\Rightarrow 160 \pm 12$

$\bar{x} \pm 3\sigma$

$\Rightarrow 160 \pm 18.$

Weights.
$(X) \rightarrow N.D$

45
63
47
53
⋮

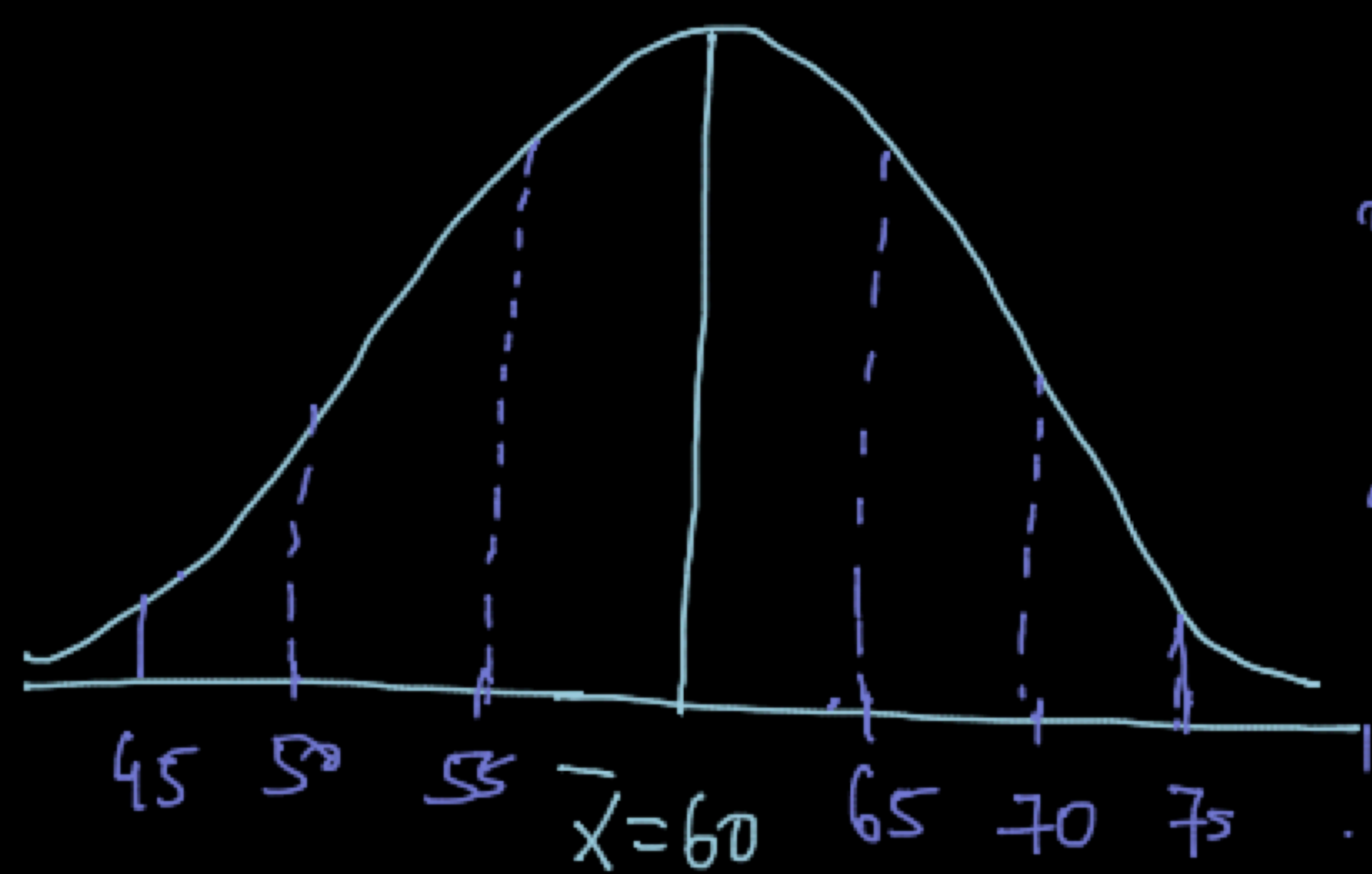
$\bar{X} = 60\,kg \quad \sigma = 5\,kg.$

$X \sim N(\mu, \sigma).$

$68\% = x \pm \sigma.$

$95\% = x \pm 2\sigma.$

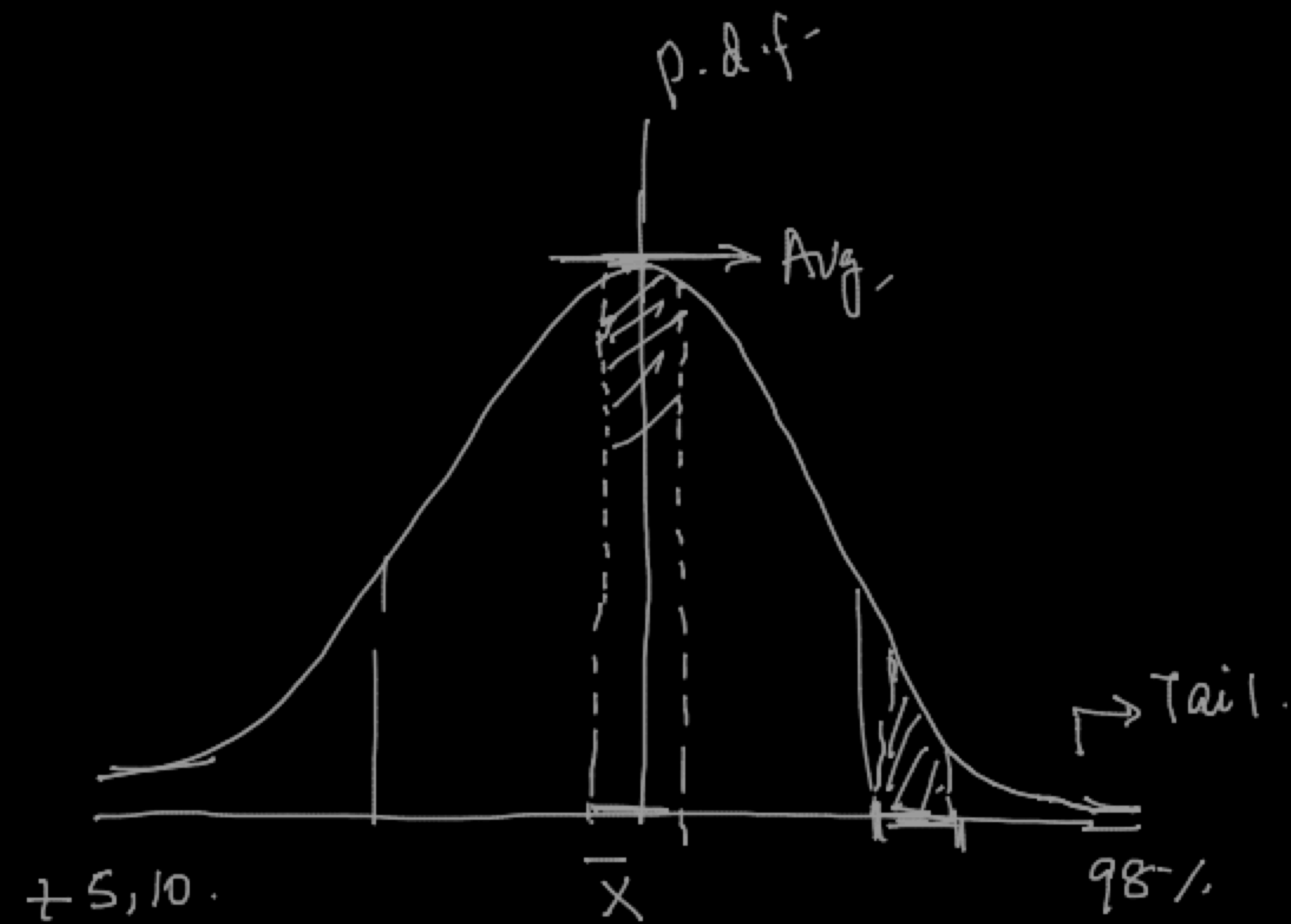
$99.7\% = x \pm 3\sigma$

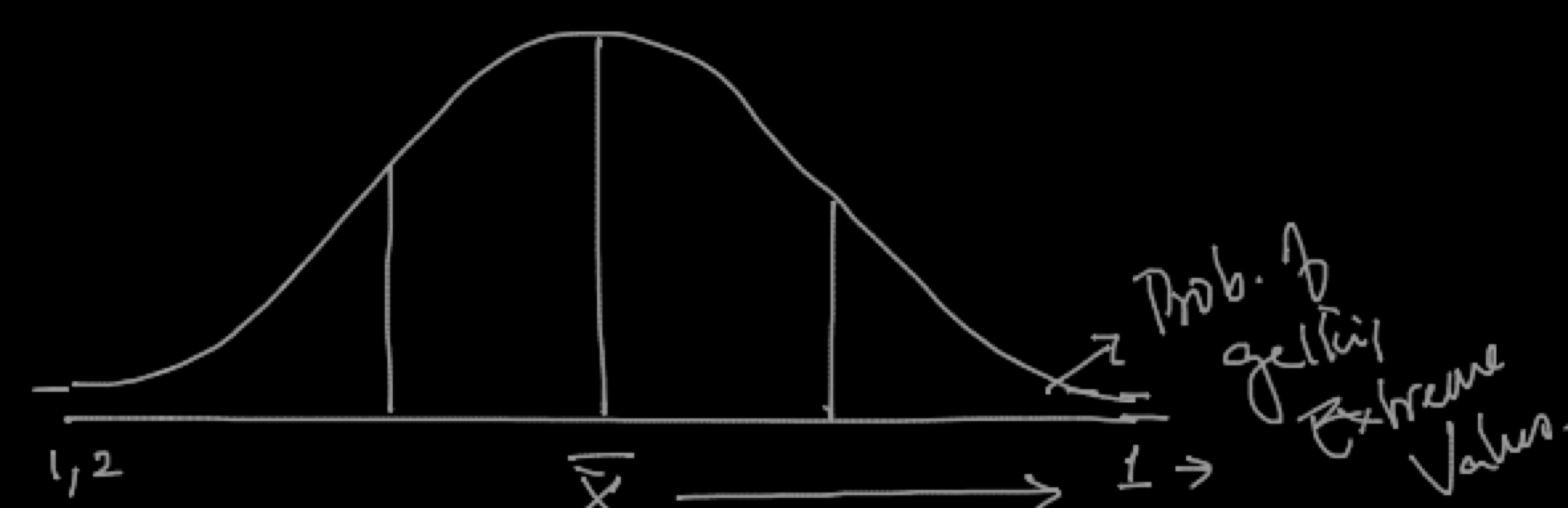


$\leftarrow 68\% \rightarrow$

$\leftarrow 95\% \longrightarrow$

$\leftarrow 99.7 \longrightarrow$


p.d.f.
Avg.
Tail.
+5,10.
$\bar{X}$
98%.


Prob. of getting Extreme Values.
1,2
$\bar{X}$
1

A new page has been created.

15 / 15

A new page has been created.

15 / 15

	Population	Sample.
No. of datapoints	N	n
Mean	μ	$\bar{x}$
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$
Std. Deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$	$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$

S1

$n = 1000$

$\bar{x} =$

$s =$

underestimates

$N = 1.3$ Billion

$\mu = \$2000$

σ

$(\bar{x})$

Avg $= 3$

$\rightarrow x_1 \quad 3$

$x_2 \quad 3$

$\boxed{x_3 = 3}$

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

$\hookrightarrow (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2$

$\hookrightarrow$ Degree of Freedom

— No. of data points free to vary.

$\boxed{\text{Total} = 100}$ $\Rightarrow$ Constraint. $\bar{x} = ?$

$P_1 \rightarrow$

$P_2 \rightarrow$

$P_3 \rightarrow$

$\vdots$

$P_9 \rightarrow$

$\rightarrow P_{10} \rightarrow 7$

$\Rightarrow 93$

$P_1 + P_2 + P_3 + \cdots$

$\Rightarrow \sum_{i=1}^{10} P_i$

16 / 16