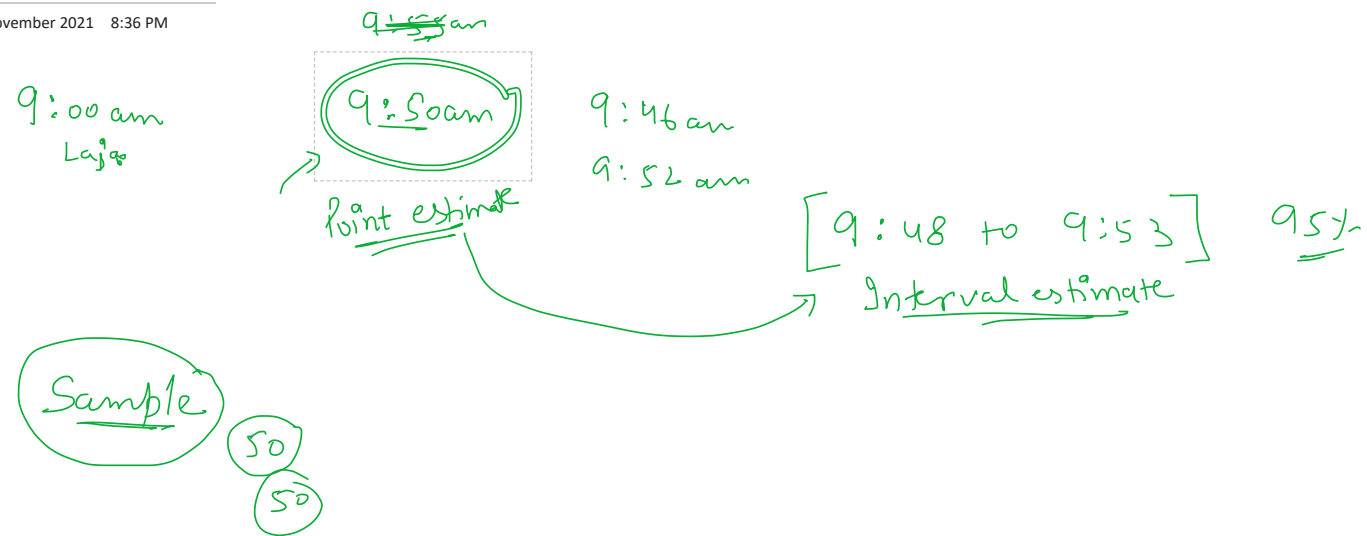


CLT_CI

Monday, 15 November 2021 8:36 PM



Q → Sample is good representative of population?

Sampling

→ Random Sampling

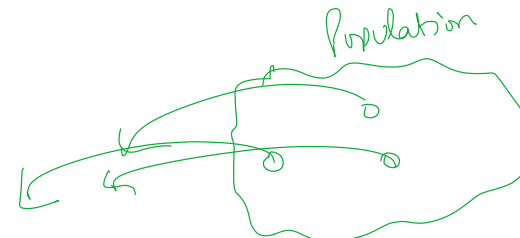
Sample Bias

Data Bias

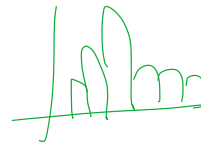
More → Male
Female

10
fair vs. Bias

Q → Size of Sample?
re-draw (resampling)



Sample data $\rightarrow [x_1, x_2, \dots, x_{50}]$



Sampling distribution

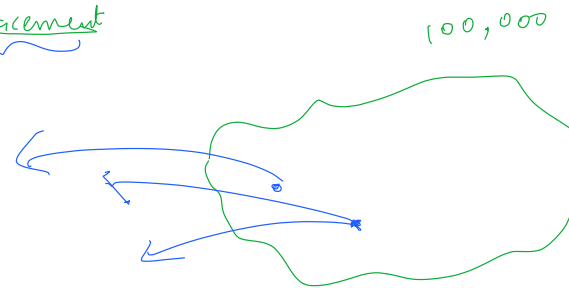
\rightarrow Variance

Bootstrapping

\rightarrow "Resampling with replacement" (random)

$n = 100$

$[x_1, x_2, x_3, \dots, x_1, \dots, x_{100}]$

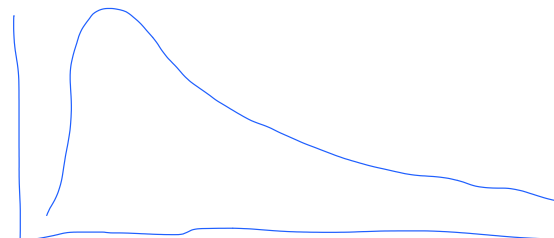


100,000

100 points

70 points

CLT [Central limit theorem]



any distribution
 $X \sim F(\text{---})$
 \rightarrow 100 datapoints

$X \sim \underline{\mu}, \sigma$

$X \rightarrow$ Bootstrapping $n=30$
 $m=1000$

- (1) Create a Sample (size = n) with replacement $S_1 = \bar{x}_1$
- (2) _____ do _____ $S_2 = \bar{x}_2$
- (3) _____ do _____ $S_3 = \bar{x}_3$
- ⋮
- ⋮
- (m) _____ do _____ $S_m = \bar{x}_m$

$$\bar{X} \Rightarrow \{ \bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_{1000} \}$$

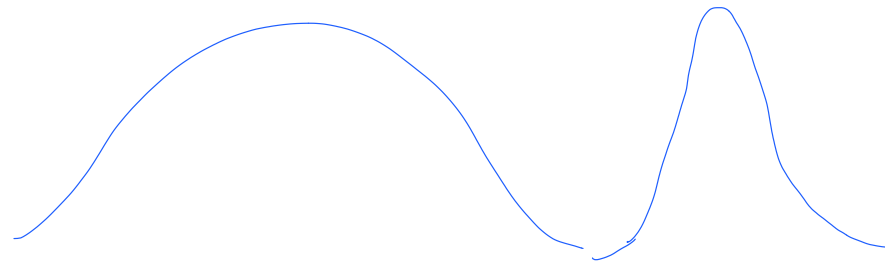
Sampling distribution of Sample mean

$$\bar{X} \sim N\left(\underbrace{\mu}_{\uparrow}, \frac{\sigma}{\sqrt{n}}\right) \quad n \rightarrow \infty$$

CLT (Central Limit Theorem)

Standard error = $\frac{\sigma}{\sqrt{n}}$
 Std of Sampling dist

OneNote
 Square root of n



95% → [x₁, x₂]

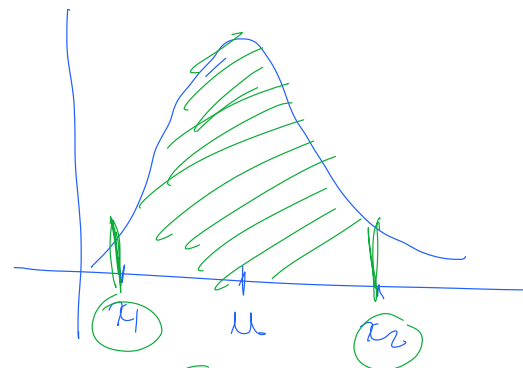
x% [5, 15]

68-95-99 rule

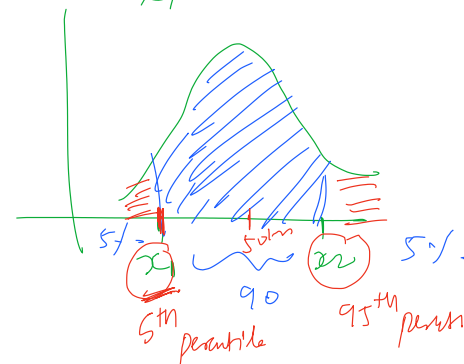
90% Confidence interval

15% ← 85% confidence interval

(90%)



$\left[\frac{\mu - 2\sigma}{x_1} \text{ to } \frac{\mu + 2\sigma}{x_2} \right]$



$$\left(\frac{100-x}{2}\right), (100-x)$$

7.5m

42.