STA 1013 : Statistics through Examples

Lecture 18: Central Limit Theorem 2

Hwiyoung Lee

October 14, 2019

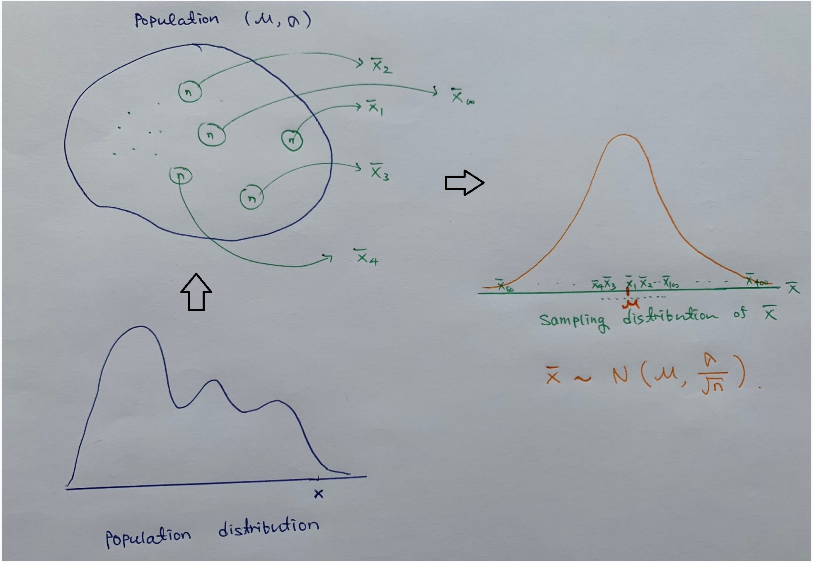
Department of Statistics, Florida State University

Review of CLT

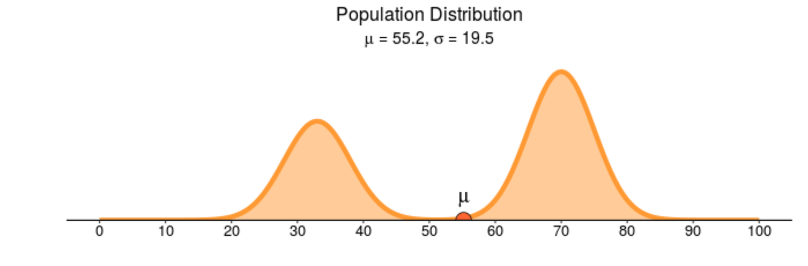
- The central limit theorem states that if you have a population with mean μ and standard deviation σ and take **sufficiently large random samples** $(n \geqslant 30)$ from the population, then the distribution of the sample means will be **approximately normally distributed**.

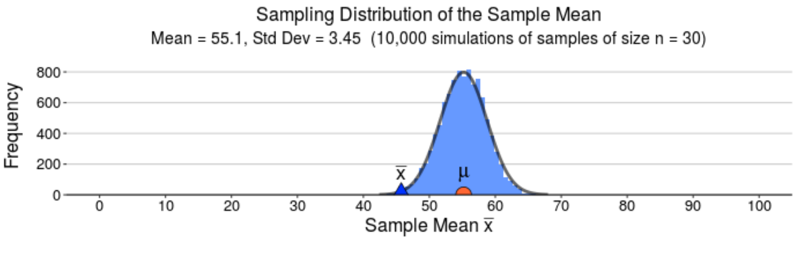
- This will hold true **regardless of** the shape of population distributions (Bell, skew, bimodalm uniform)

Review of CLT

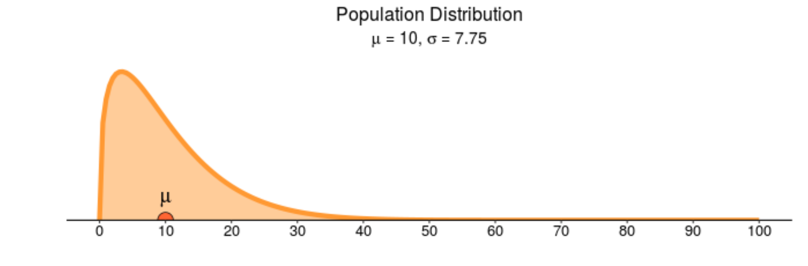


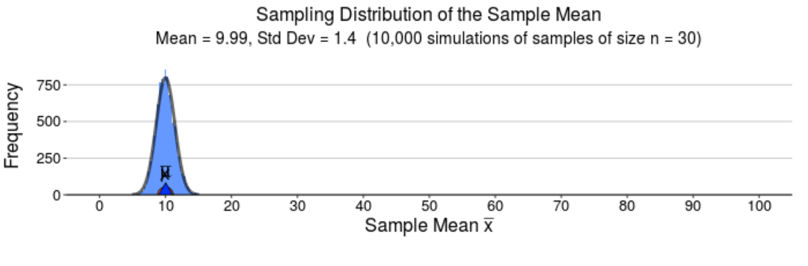
CLT Simulation (Bimodal)



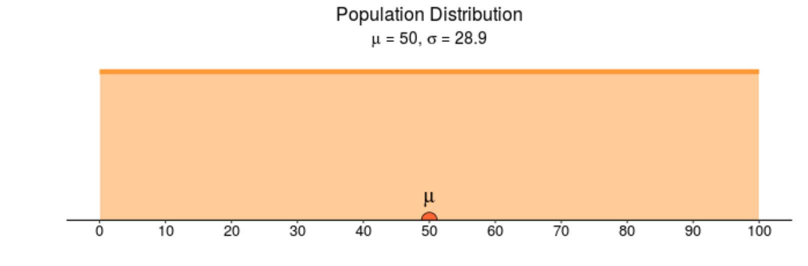


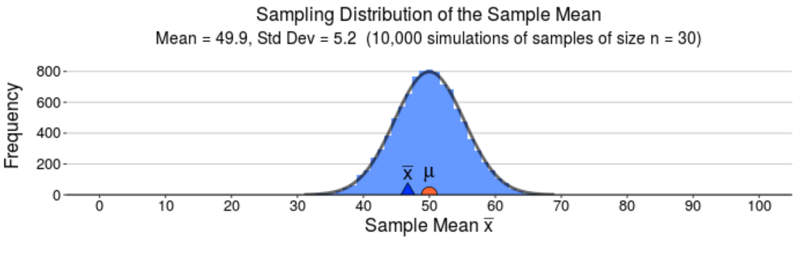
CLT Simulation (Skewed)





CLT Simulation (Uniform)





Example : CLT

- Original population of data : 2, 4, 6, 8
- Population mean $\mu = (2 + 4 + 6 + 8)/4 = 5$
- Draw random samples (size n=2)

Samples	$\bar{X}$	Samples	$\bar{X}$	Samples	$\bar{X}$	Samples	$\bar{X}$
2,2		4,2		6,2		8,2	
2,4		4,4		6,4		8,4	
2,6		4,6		6,6		8,6	
2,8		4,8		6,8		8,8	

Average of the sample means : ____________

Example : CLT

Distribution of $\bar{X}$

2	3	4	5	6	7	8

Sampling distributio of $\bar{X}$ roughly follows normal distribution

Example : CLT

1. How many keys do you have on you right now? ___________

2. Fill in the table below

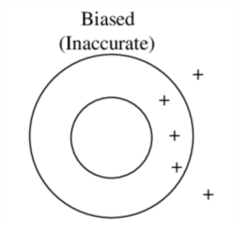
Group Member Name	Number of keys
	A =
	B =
	C =
	D =

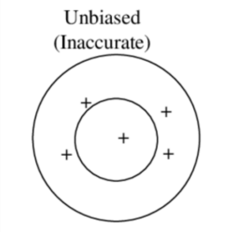
3. Calculate the following averages

(A+B)/2		(A+C)/2		(A+D)/2	
(B+C)/2		(B+D)/2		(C+D)/2	

Central Limit Theorem

- $\bar{X}$ is an estimator of the population mean μ because we can, and we often do, use our particular sample result $\bar{X}$ as an estimate of μ.
- $\bar{X}$ is an **unbiased estimator** of the population mean, because the mean of the $\bar{X}$'s is equal to the population mean.

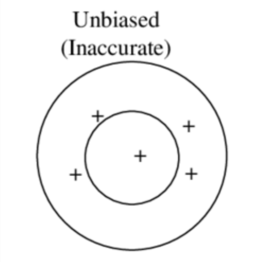


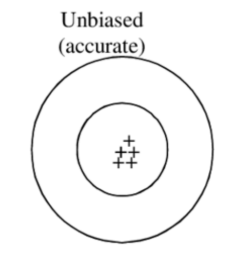


Central Limit Theorem

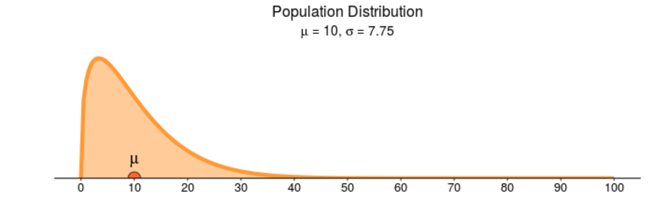
- In general, sample means vary less than individual values.
  - Individual value : $X \sim N(\mu, \sigma)$
  - Sample means : $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

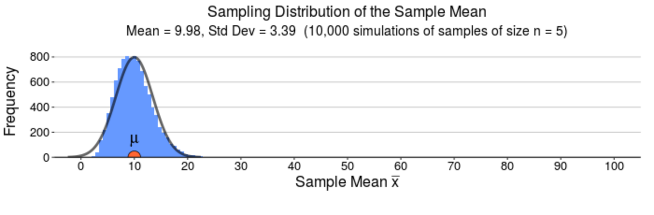
- In general, larger samples produce values of $\bar{X}$ that stay closer to μ. It follows that a **larger sample is more likely to give us a better estimate of μ than a smaller sample**.



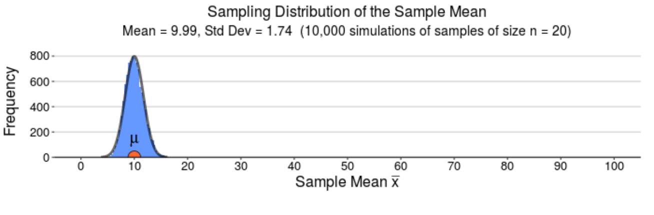


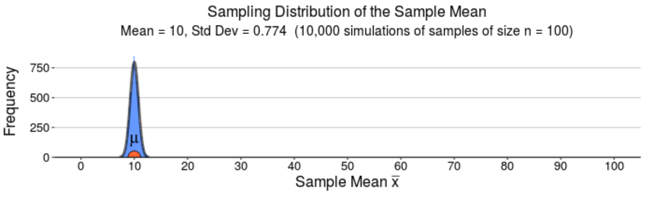
Central Limit Theorem





Central Limit Theorem





Exercise

Male weight

Assume that weights of men are normally distributed with a mean of 170 lb and a standard deviation of 30 lb.

- What percentage of **individual** men have weights less than 185 lb?

- If samples of 36 men are randomly selected and the mean weight is computed for each sample, what percentage of **the sample means** are less than 185 lb?

Male weight

- What percentage of **individual** men have weights greater than 167 lb?

- If samples of 100 men are randomly selected and the mean weight is computed for each sample, what percentage of **the sample means** are greater than 167 lb?

Test Scores

You are a middle school principal and your 100 eighth-graders are about to take a national standardized test. The test is designed so that the mean score is $\mu = 400$ with a standard deviation of $\sigma = 70$. Assume the scores are normally distributed.

- What is the likelihood that one of your eighth-graders, selected at random, will score below 375 on the exam?

Test Scores

- Your performance as a principal depends on how well your entire group of eighth-graders scores on the exam. What is the likelihood that your group of 100 eighth-graders will have a mean score below 375?

Test Scores

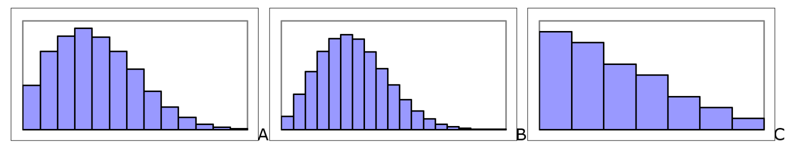
Two sampling distributions of $\bar{X}$ are described below. Which sampling distribution will have the lower standard deviation?

1. SRSs of 4 subjects each are selected and the mean is calculated for each sample.

2. SRSs of 16 subjects each are selected and the mean is calculated for each sample.

Sample Size

Match the sampling distributions of $\bar{X}$ shown below to these sample sizes: n=1, n=2, n=3.



When n=1, the sampling distribution of $\bar{X}$ is the same as the distribution of the individual values X, so it is the same as the original population distribution.

LSQA

Section 15 : MWF (10:10 $\sim$ 11:00)

- Test Date : Oct 21 (Mon) HCB 217

Section 05 : MWF (12:20 $\sim$ 1:10)

- Test Date : Oct 18 (Fri) OSB 108