

## 7.4 & 7.5 Sampling Distributions and the Central Limit Theorem

Parameters, Statistics, and inferences

A statistic is a numerical measure describing a sample.

Notation of Statistics and Parameters		
Measure	Statistic	Parameter
Mean	$\bar{x}$ (x-bar)	$\mu$ (mu)
Variance	$s^2$	$\sigma^2$ (sigma squared)
Standard Deviation	$s$	$\sigma$ (sigma)
Proportion	$\hat{p}$ (p-hat)	$p$

### Inferences

is the logical process of reaching a conclusion based on evidence, reasoning, and established facts,

Ex: if there is smoke, there will be fire. we are not 100% sure.

### Types of Inferences

# Types of Inferences

1. Estimation: we estimate the value of a population parameter using a sample
2. Testing: we do a test to help us make a decision about a population parameter
3. Regression: we make predictions or forecasts about a statistic

**Requires understanding sampling distributions and the Central Limit Theorem**

## Frequency vs. Sampling Distribution

### Frequency Distribution

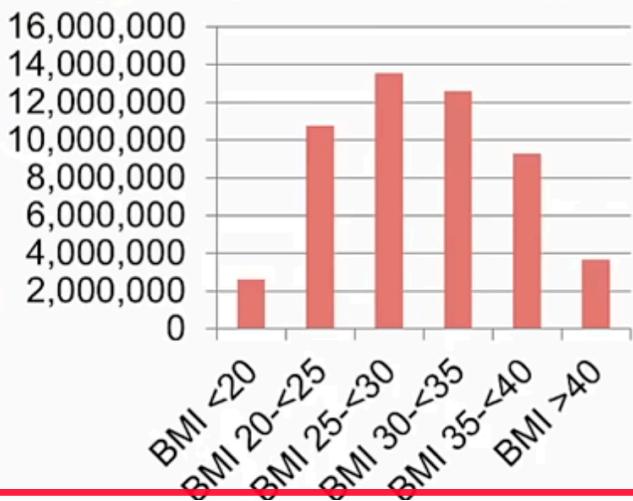
1. Make a histogram of a quantitative variable.
2. Draw the shape and name the distribution.



### Sampling Distribution

1. Start with a population.
2. Decide on an n.
3. Take as many samples of n as possible from the population.
4. Make an  $\bar{x}$  for each sample.
5. Make a histogram of all the  $\bar{x}$ -bars.

# Definition of a Sampling Distribution

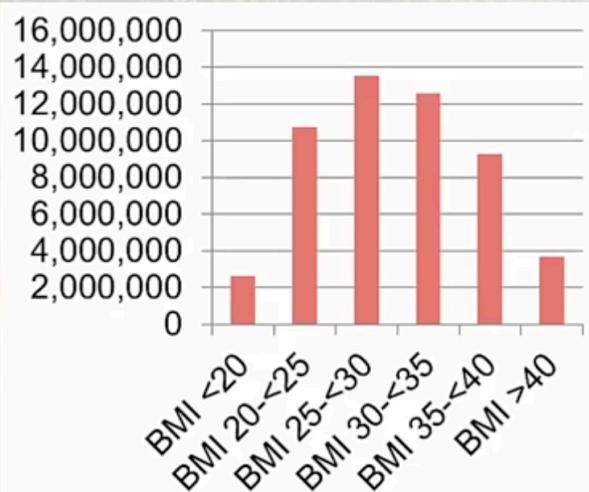


- A sampling distribution is a probability distribution of a sample statistic based on all possible simple random samples of the same size from the same population.

Central Limit Theorem

For any normal distribution:

1. The sampling distribution (the distributions of x-bars from all possible samples) is also a normal distribution
2. The mean of the x-bars is actually  $\mu$
3. The standard deviation of the x-bars is actually  $\sigma/\sqrt{n}$



# Central Limit Theorem: In Formulas

$$\mu_{x\text{-bars}} = \mu$$

$$\sigma_{x\text{-bars}} = \sigma / \sqrt{n}$$

$$Z = \frac{x\text{-bar} - \mu_{x\text{-bars}}}{\sigma_{x\text{-bars}}} = \frac{x\text{-bar} - \mu}{\sigma / \sqrt{n}}$$

....where

- n is the sample size
  - Note: Only works when  $n \geq 30$ !
- $\mu$  is the mean of the x distribution (population mean), and
- $\sigma$  is the standard deviation of the x distribution (population standard deviation)

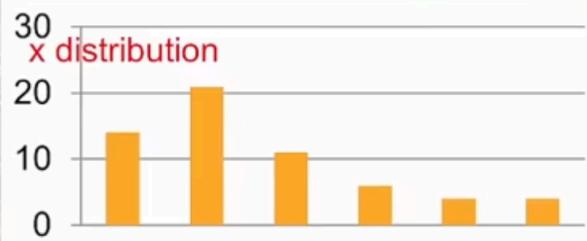
The standard error is the standard deviation of the sampling distribution.

For the  $x\text{-bar}$  sampling distribution, standard error (SE) =  $\sigma / \sqrt{n}$

Standard Error is the standard deviation of the sampling distribution.

## Central Limit Theorem

1. If the distribution of x is normal, then the distribution of  $x\text{-bar}$  is also normal.
2. Even if the distribution of x is NOT normal, as long as  $n \geq 30$ , the Central Limit Theorem says that the  $x\text{-bar}$  distribution is approximately normal.



# What Are We Doing?

## Chapter 7.1-7.3

- We had a normally distributed  $x$ .
- We had a  $\mu$  and a  $\sigma$ .
- We want to find the probability of selecting a value (from the population) above or below a value of  $x$ , so we use the z-score and z-table for probabilities.
- We used this formula: 
$$z = \frac{x - \mu}{\sigma}$$

## Chapter 7.4-7.5

- We have a normally distributed  $x$ .
- We have a  $\mu$  and a  $\sigma$ .
- We want to find the probability of selecting a sample  $n$  (from the population) with a mean value ( $x\bar{}$ ) above or below a value of  $x\bar{}$ , so we use the z-score and z-table for probabilities.

We will use this formula: 
$$z = \frac{x\bar{}}{\sigma/\sqrt{n}}$$

## What $\mu$ (mu) is

- $\mu$  = population mean
- It's the true average of everyone in the population
- Fixed number (but usually unknown)

👉 Example:

Average height of all students in a university =  $\mu$

## What ( $x\bar{}$ ) is

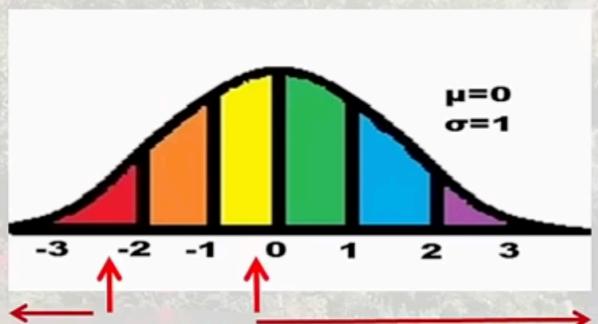
- $x\bar{}$  = sample mean
- Calculated from a sample, not everyone
- Changes from sample to sample (random)

👉 Example:

Average height of 30 students you randomly pick =  $x\bar{}$

# Remember the Students?

- Assume the 100-student class is a population.
- Now I have to pick an n
  - Let's pick 36.
- Question: What is the probability of me selecting a sample of 36 students with an  $\bar{x}$ -bar between 60 and 65?



$$z1 = (60-65.5)/2.4 = -2.28$$

$$p1 = 0.0113$$

$$z2 = (65-65.5)/2.4 = -0.21$$

$$1 - 0.0113 - 0.5832 = 0.4055 \quad p2 = 0.5832$$