* **Parameters & Statistics**
* Inference: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ characteristics of a \_\_\_\_\_\_\_\_\_\_\_\_\_ using \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
* A **parameter** is a number that describes some \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of the **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
* The value of a population characteristic is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* A **statistic** is a number that describes some characteristic of a **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
* Different samples 🡪 different \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* Practice:

A business owner wants to know the average household income in her business’s zip code. She randomly samples 80 households in the zip code and finds their mean income to be $46,144.

* What is the parameter in this case:
* Statistic:
* **Sample Distribution & Sampling Distribution**
* Sample distribution is the distribution of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* Sampling distribution:
* Fix the sample size and select samples from a population
* Compute statistics from sample data
* Sampling distribution is the distribution of **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
* **Example:**

Identical balls in a black box with different numbers 1, 2, and 3. What is the distribution of the sample mean with a sample size of 2?

* **Accuracy (bias) vs. precision (variation) ---- How sure are we of our ­­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_?**
* How to make it accurate?
* How to reduce variability?
* **Sampling Distribution for a Mean**
* Population Mean：
* Estimator:
* ,

1. We are centered at \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, meaning our estimates are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
2. Higher sample size means **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
3. The higher the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ in the population data, the higher the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ in the sampling distribution

* **Central Limit Theorem (CLT)**

If you have a population with mean μ and take **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** samples from the population **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**, then the distribution of **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** will be **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.**

**Conditions:**

1. the sample was **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** from the **target population** (in the case of an experiment, you must ensure there was random assignment to treatment)

If condition is not satisfied, the estimator is **biased**. If it is satisfied, the sampling distribution is **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

1. 10% Condition

The **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** must be less than \_\_\_\_\_\_\_\_\_\_ of the **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

This condition ensures that: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

For example if n = N, = \_\_\_\_**\_\_\_\_\_\_\_\_\_\_\_**

1. Large sample size: the sample size is **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

This provides evidence that the sampling distribution is **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

* **Practice 1：**

Every day people watch 1 billion hours of videos on YouTube. That breaks down to every single person on earth watching YouTube videos for about 8.4 minutes per day. For U.S. teens, in any given day, the amount of time spent watching YouTube videos is approximately Normal with mean 18.5 minutes and standard deviation 5.3 minutes.

a. Find the probability that a randomly chosen U.S. teen watches YouTube for more than 25 minutes in a given day.

b. Suppose we choose an SRS of 10 U.S. teens. Let = the mean amount of time spent watching YouTube videos for the sample. What is the mean of the sampling distribution of ?

1. Calculate and interpret the standard deviation of the sampling distribution of . Verify that the 10% condition is met.
2. Find the probability that the mean amount of time spent watching YouTube for the teens in the sample exceeds 25 minutes.

* **Practice 2：**

Among iPhone users who share their data with Apple, the weekly screen time is skewed to the right with µ = 13.5 hours and σ = 3.75 hours. A random sample of 100 iPhone users are selected and the mean weekly screen time () of the sample is calculated.

1. Describe the shape of the sampling distribution of for samples of 100 randomly selected iPhone users.
2. Find the mean and standard deviation of the sampling distribution of . Be sure to check the 10% condition.
3. Calculate the probability that the weekly screen time for the sample is between 12 and 13 hours.