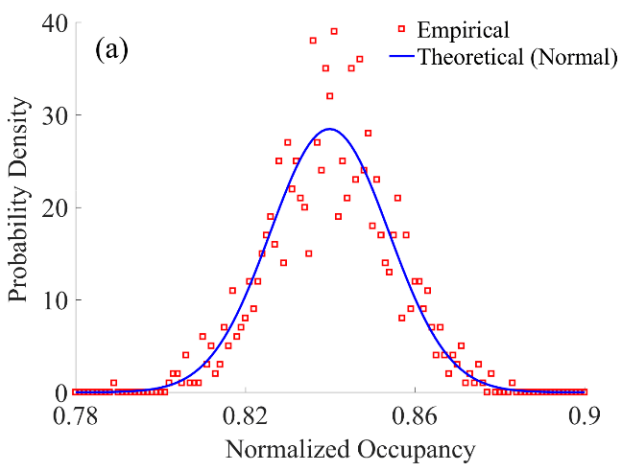
Population, Sample, and Sampling Distributions

Empirical Distributions and Theoretical Distributions

Empirical is a distribution from a random sample used for the estimation of a true distribution.

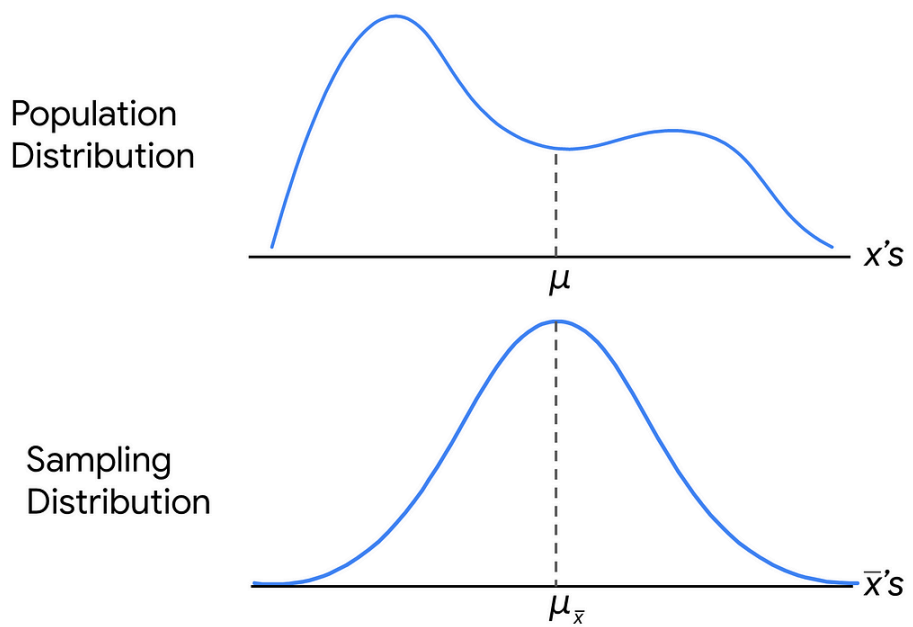
Theoretical is a distribution received by a set of logical and mathematical reasoning from given principles or assumptions.



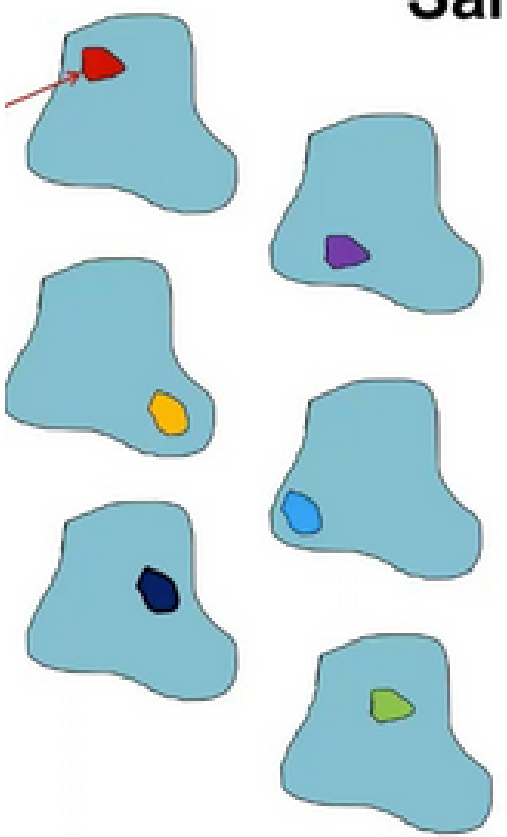
Population Distributions and Sample Distributions

Population Distributions

Sample Distributions

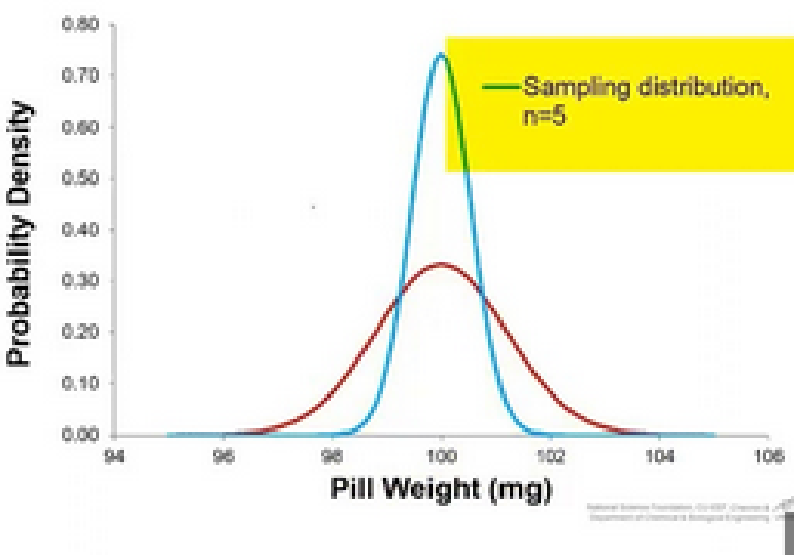


Population Distributions and Sample Distributions



Sampling Distribution

1. Take sample of size n
2. Compute $\bar{X}$ of those n items
3. Repeat many times to generate a distribution of $\bar{X}$



Distribution Terms

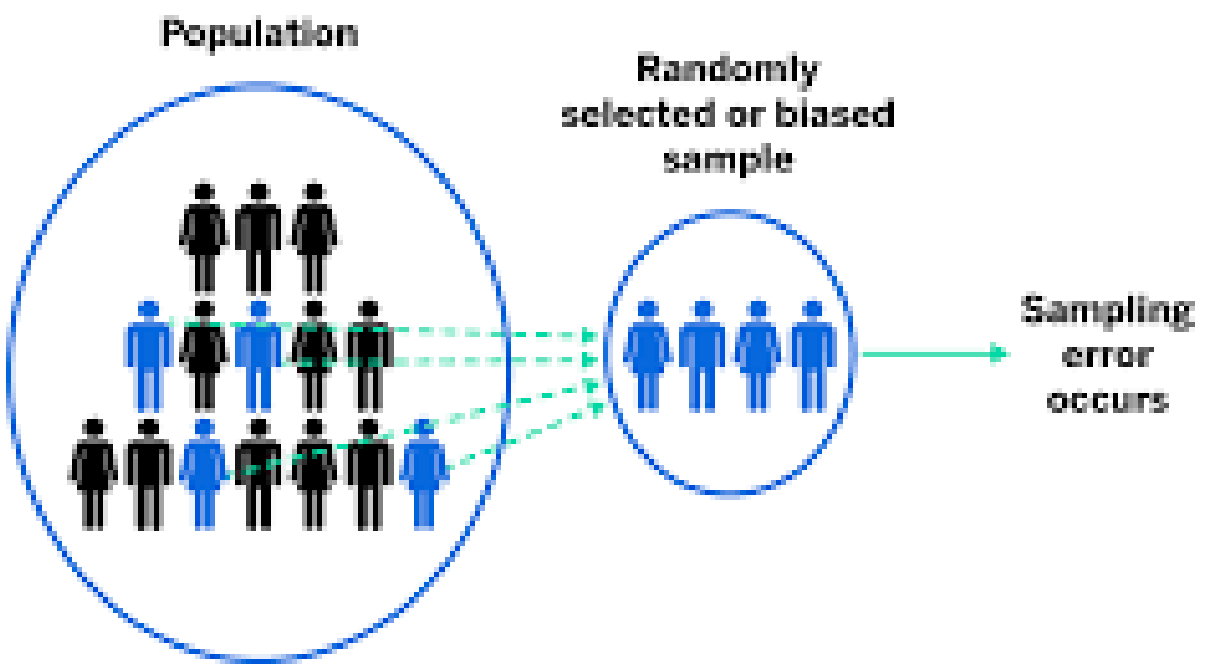
Representativeness is how closely the characteristics of a sample match the population

Statistic is a numerical characteristic or measure that describes a sample.

Parameter is a variable that appears in an **equation or function** but remains **constant** within a specific context. For example, means or standard deviations.

Sampling Error

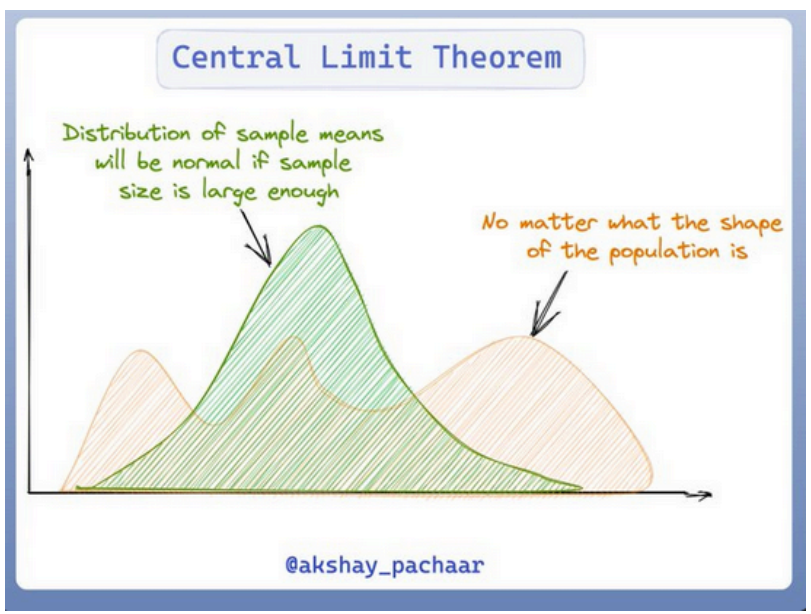
The difference between a sample statistic (such as the sample mean) and the corresponding population parameter (such as the population mean).

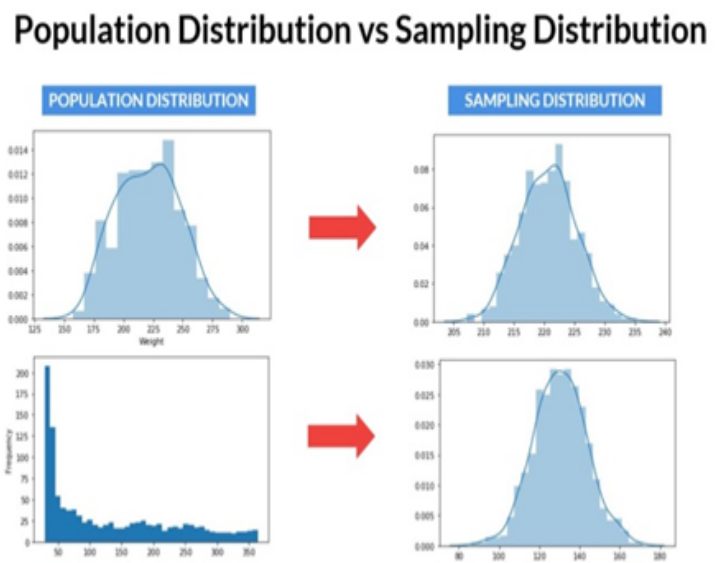


Central Limit Theorom

States that the sampling distribution of the sample mean approaches a normal distribution as the sample size increases, regardless of the shape of the population distribution, provided that the sample size is sufficiently large.

Let's look at another fun visual here





Standard Error

A measure of how much the sample statistic is expected to vary from one sample to another.

A high standard error indicates a less precise estimate, which means that if you were to repeat the study and draw another sample, the results (like the mean) might be further away from the population mean on average.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

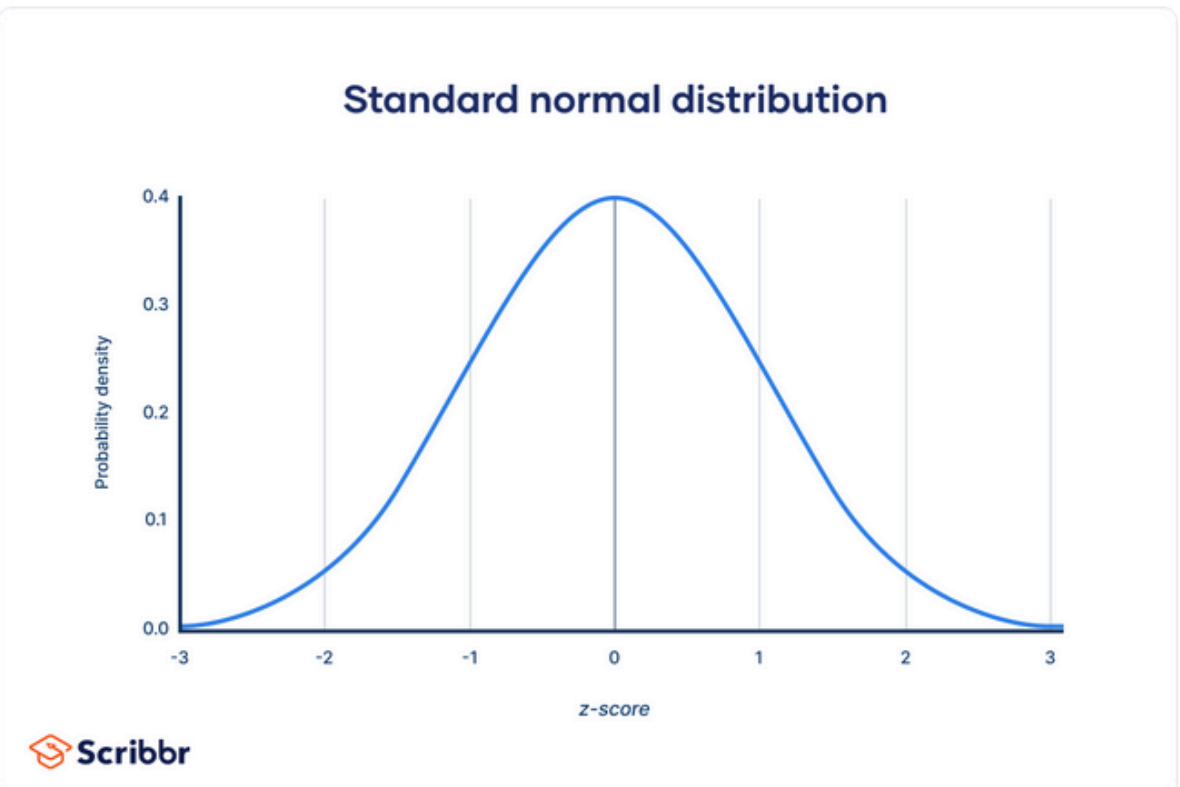
s = standard deviation

n = the sample size

The z Distribution

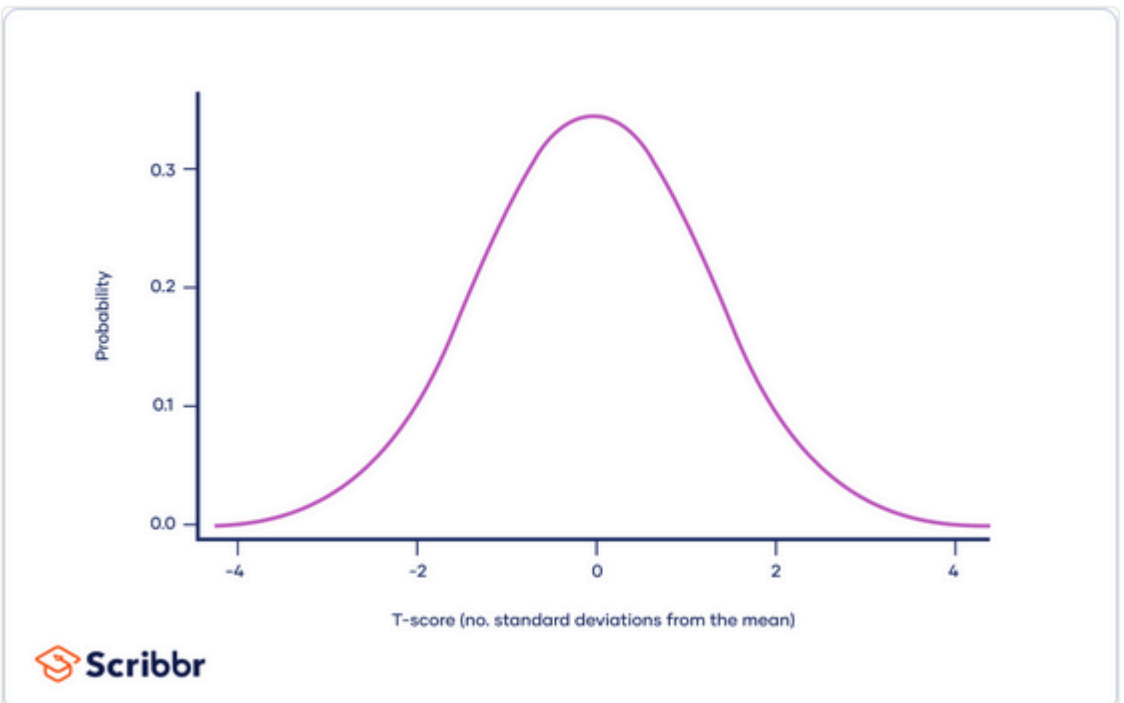
The z-distribution, (i.e., a standard normal distribution), is characterized by a bell-shaped curve that is symmetric around its mean.

Used when we have a large sample size.



The t Distribution

Is a probability distribution that arises in statistics when estimating the population mean from a sample with a small sample size or when the population standard deviation is unknown.



Have a great day!