

[← Go Back to Foundations for Data Science](#)

Course Content

2.1.11 Summary - Sampling and Inference Foundations-2

In the real world, it is not always possible to get an exact idea about the parameters of the population. This is why it becomes important to take samples from the population. The idea behind inferential statistics is to approximate the population parameters by taking samples from it. Hence, a good sample is one that represents the population well.

To understand this, let us consider a population of 1000 people, where we want to approximate their average age. First, we will extract 5 samples of size 20 from the population, i.e., each sample contains 20 people.

Now, let \bar{x} be the random variable that represents the sample mean of each sample and let the sampling distribution of \bar{x} be [23, 33, 12, 24, 35], where 23 represents the average age of the first sample of 20 people, 33 represents the average age of the second sample of 20 people, and so on.

We can use this sampling distribution to approximate the **unknown population parameters** like the mean (denoted by the Greek letter μ) and the standard deviation (denoted by the Greek letter σ) using the following rule:

The mean of the sampling distribution of \bar{x} is μ and the standard deviation of the sampling distribution of \bar{x} is σ/\sqrt{n} .

Note: The above rule is based on one of the most famous theorems in statistics, known as the **Central Limit Theorem (CLT)**. The next couple of video lectures will help you get clarity on CLT as well as the above rule.

<https://bcheggeseth.github.io/Stat155Notes/sampling-distribution-and-clt.html>

You can read more about sampling distributions [here](#).

Happy Learning!

[< Previous](#)

[Next >](#)

Proprietary content.©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

© 2023 All rights reserved.

[Help](#)