

Central Limit Theorem and Confidence Intervals - Introduction

Introduction

In this section, you'll be introduced to inferential statistics. You'll learn about sampling, the central limit theorem, and the T-distribution.

Distributions and Sampling

In this section, we're returning to statistics to broaden and deepen our understanding of distributions and sampling.

Sampling

We'll start by providing an introduction to the idea of **Sampling** - selecting a subset of a population to survey. We'll then start to introduce some statistics related to sampling by explaining and showing how to calculate the standard error.

The Central Limit Theorem

Once we understand a bit about sampling, we'll explore how we can use it by digging deep into one of the coolest and most important concepts in inferential statistics--the **Central Limit Theorem**! We'll start by learning about how the Central Limit Theorem works, and explore how we can use it in a way that allows us to treat non-normal distributions as normal distributions, and provides a way for us to estimate parameters about a population.

The T-Distribution

Finally, we'll end this section by learning about how we can use the **T-Distribution** for dealing with samples that are smaller, and that have an unknown standard deviation. We'll explore how the T-Distribution works, learn about *degrees of freedom*, and then see how we can calculate confidence intervals using our newfound knowledge of the T-Distribution.

Summary

While some of this material may seem a little dry, a deep understanding of and intuition for distributions and sampling will be important in your career as a data scientist. This knowledge will help you avoid making mistakes in your EDA (exploratory data analysis), feature selection, and modeling work, which could lead to faulty predictions from your models.