

Statistical Distributions - Introduction



[_ \(https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2\)](https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2)  [_ \(https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2/issues/new/choose\)](https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2/issues/new/choose)

Introduction

This is the second of two sections covering foundational statistical concepts. Now that you have covered the essential set theory and probability theory, you'll learn more about the formal ways to represent statistical distributions.

Descriptive to Inferential Statistics

Previously in the course, you learned about the *descriptive* statistics used to describe a distribution of data — in particular, measures of centrality (e.g. mean, median) and measures of spread (e.g. variance, standard deviation).

In most real-world data science contexts, you will not have access to complete information about a distribution of data. Instead, you will have a **sample**. In order to make claims about the complete population of data, you'll need to perform **inference** (i.e. "inferential statistics") using the available sample data.

Statistical Distributions

In order to understand how to make these inferences, first you'll need some additional understanding of different kinds of distributions, how they relate to the underlying data types being represented (discrete vs. continuous), and how we represent them formally using the math notation introduced in the previous section.

In particular, we'll look at ways of representing **probability distributions** using the **Probability Mass Function** (for discrete data) and **Probability Density Function** (for continuous data), as well as another statistical distribution represented by the **Cumulative Distribution Function**.

We'll also dig into some of the specific distributions that data points often fall into, including the Binomial and Bernoulli distributions (for discrete data) and the Normal distribution (for continuous data). We'll conclude by introducing the concepts of skewness and kurtosis, which help to quantify how "un-normal" a given distribution is.

Appendix

There are far more statistical distributions than we have time to cover in this course. If you are interested in digging deeper into distributions, you can find additional content in the Appendix.

Summary

In this section, we're going to connect the descriptive statistics concepts from earlier in the course to the set theory and probability theory from the previous section, to provide a foundation for inferential statistics.

How do you feel about this lesson?



Have specific feedback?

[Tell us here!](https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2/issues/new/choose)  [\(https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2/issues/new/choose\)](https://github.com/learn-co-curriculum/dsc-distributions-intro-v2-2/issues/new/choose)