

Introduction to Linear Regression - Introduction



<https://github.com/learn-co-curriculum/dsc-linear-regression-section-intro-v2-5>



<https://github.com/learn-co-curriculum/dsc-linear-regression-section-intro-v2-5/issues/new/choose>

Introduction

In this section you will apply your statistics and hypothesis testing knowledge and create your first statistical model: a simple linear regression! Most common statistical tests are special cases of linear models, and linear models are the conceptual bridge between traditional statistics and machine learning.

Statistical Modeling

First we'll set the stage by discussing how all statistical tests are based on underlying mathematical assumptions, and how we apply them to specific contexts. In particular we'll focus on *linear models*.

Linear Regression

The most basic linear model is a *simple linear regression*, which looks like a basic $y = mx + b$ best-fit line.

Regression Model Evaluation

All models are wrong, but some are useful. ([George Box](#))

https://en.wikipedia.org/wiki/All_models_are_wrong

Like with other statistical tests, you can evaluate whether or not a simple linear regression model is statistically significant using p-values.

In addition to questions of statistical significance, we can also measure regression model performance in terms of *goodness of fit* and *error*.

Linear Regression in StatsModels

The library StatsModels contains high-quality tools for building and evaluating linear regression models. We'll practice setting up these models and interpreting their results.


Summary

Congratulations, you've made it to statistical modeling! This portion of the course can be tough conceptually, but it underpins all kinds of statistical and machine learning approaches.

How do you feel about this lesson?



Have specific feedback?

Tell us here!  <https://github.com/learn-co-curriculum/dsc-linear-regression-section-intro-v2-5/issues/new/choose>