ESMA 6205: Introduction Linear Models

Alejandro Ouslan

Spring 2025

Contents

1 Introduction to the Scientific Method

Process by which scientist, collectively and over time, endeavor to construct an accurate representation of the world.

- Observation and description of a phenomenom or group of phenomena.
- Formulation of an hypothesis to explain the phenomena.
- Use of the hypothesis to predict the existence of other phenomna, or to predict quantitatively the results of new observation
- Performance of experimental tests of the predictions by several independent experimenters and properly performed experiments.

2 What is statistics?

- Statistics science of learning from data, and of measuring, controlling, and communication uncertainty.
- Provides the navigation essential for controlling, and societal advances

3 Why become an statistician/data sciences?

- "The best thing about being a statistician is that you get to play in everyone else's backyard", John Tukey, Princeton University.
- Data science combines math and statistics, specialized programing, advanced analytics, artificial intelligence (AI) and machine learning with specific subject matter expertise to uncover insights in data.
- Statiscian apply statistical thinking and methods to a wide variety of scientific, social, and buisness endeavors in such areas as astronomy, biology, education, economics, engineering, genetics, marketing , medicine, psychology, public health sports, among many.

4 Paths in Stats

Like many choices in life, frequentist and bayesian.

- **Statistics Theory:** provides an underlyin rational and provides a consistent basis for the choice of methodology.
- **Applied Statistics:** comprises descriptive statisics and the application of inferential statistics.

5 Notation Notation and Simple Matrix Algebra

- n: number of distinct data points, or observvations, in our sample.
- p: number of distinct that are available for use.
- $\mathbf{X}$: matrix
- $\mathbf{y}$: Vector

6 Why linear models?

- Linear models are a fundamentals in statistics, data science and machine learning to understand relationships between variables.
- Linear

item