DS 5740 | Advanced Statistics

I. Course Information

Instructor: Alexander Christensen, Ph.D. alexander.christensen@vanderbilt.edu

Class Times: 9:00-10:15am on Mondays and Wednesdays

Classroom: Sony Building 2001-A

Office Hours: 10:30-11:30am Monday; by appointment

Learning Management System: Brightspace

Description: Predicting future outcomes is fundamental to human activity. Understanding how past events lead to current conditions can enables us to make predictions about the future. This course will focus on how we can make better predictions about the future based on past evidence. Applied uses include forecasting market outcomes (e.g., stocks, company earnings) to human behavior (e.g., planning your mortgage, public health policy).

The majority of the course covers time series forecasting to use past performance patterns and make predictions about the future. Key topics in this component includes graphical representations, decomposition, smoothing, ARIMA-type models, and dynamic covariation. A second component of this course will focus on A/B testing (also referred to as experimental design). In this component, we'll cover how we can predict outcomes by manipulating different features of the environment. Finally, depending on time, we'll briefly cover Bayesian modeling. For this component, we'll review Bayes Theorem and explore some applications of Bayesian models, which use past information and current evidence to update our understanding of relationships between variables.

This course will use R statistical software and packages. All examples in the course will be provided through R code and all modeling can be performed using open-source R packages. Accompanying books are free and based in R. We will not cover all of the material in any one book, but they are valuable resources for you to hone your skills and continue your career path.

Goals of the Course:

1. Understand when forecasting can be applied and what model is most appropriate

2. Use R to manipulate data and perform forecasting analyses using popular packages such as {tsibble}, {fable}, and {prophet}

3. Predict real-world outcomes

4. Know how to construct experimental designs that can identify ways to improved performance

Tentative Course Schedule

Week 1 (08.24): Introduction to course

- Psychology of teams, A/B testing example, download R software and packages, first assignment (pick a stock)

Time Series Forecasting

Week 2 (08.29 & 08.31): Introduction to time series, {tsibble}, and plotting

Week 3 (09.05 & 09.07): Time series decomposition and features

Week 4 (09.12 & 09.14): Diagnostics, forecasts, and evaluation

Week 5 (09.19 & 09.21): Time series regression models

Week 6 (09.26 & 09.28):

Week 7 (10.03 & 10.05):

Week 8 (10.10 & 10.12):

Fall break on 10.13 & 10.14

Week 9 (10.17 & 10.19):

Week 10 (10.24 & 10.26):

Week 11 (10.31 & 11.02):

Week 12 (11.07 & 11.09):

Week 13 (11.14 & 11.16):

No courses this week – holiday break

Week 14 (11.28 & 11.30):

Week 15 (12.05 & 12.07):

Books (all available for free):

Time series forecasting

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (3rd edition). Melbourne, AUS: OTexts. https://otexts.com/fpp3/

A/B testing

Taback, N. (2022). *Design and analysis of experiments and observational and studies using R* (1st edition). New York, NY: Chapman and Hall/CRC. http://designexptr.org/index.html

Bayesian modeling

Albert, J. (2007). *Bayesian computation with R*. New York, NY: Springer. https://bayesball.github.io/bcwr/index.html

Software

R: https://cran.r-project.org/

Rtools (Windows only): https://cran.r-project.org/bin/windows/Rtools/rtools42/rtools.html

RStudio: https://www.rstudio.com/products/rstudio/download/#download