R Companion to Discrete Choice Methods with Simulation

Dan Yavorsky

Table of contents

Welcome

> You don't understand it, until you have coded it.

It's a mantra that my mentors instilled in me in graduate school, and one that I propogate on my students today. I believe it in deeply. If you want to understand a statistical model, it is insufficient to interact with the model only through pen and paper, no matter how well you organize the subscripts or how masterfully you interweave elements of the Greek and Roman alphabets. To really understand a statistical model, you need to be able to simulate data from the model, and then take that simulated data to an estimation routine that enables you to recover the parameters of interest.

This is something you learn to do in graduate school. But no one teaches it to you! Instead, it's a skill that students develop independently and inefficiently between classes and assignment due dates. There is a tremendous imbalance between how immensely important this skill is and the lack of time and instruction dedicated to it.[1] That imbalance is the motivation for this book.

How to read this book

As the title suggests, the structure, topics, and organization of this book parallel Kenneth Train's masterful text *Discrete Choice Methods with Simulation* (Second Edition) freely available online at https://eml.berkeley.edu/books/choice2.html.

You should read — or most likely, re-read — one chapter from Train (2009) and then read and work through the corresponding chapter here, alternating between our texts. Do not simply read this book from start to finish without frequenly returning to Train (2009). Doing so only deprives you of the intended experience and likely dramatically reduces the amount you will learn from the process.

In addition, I strongly encourage you write or copy the code as you work though this book. I would mandate this if I could, but I'm not sitting there next to you and so I must settle for simply sharing my encouragement. Write the code. Play with the code. Break the code. Make

[1] What I'm talking about is the field of statistical computing. My experience was that graduate students in economics, psychology, and business had little, if any, formal exposure to these ideas. For a related text that emphasizes computational statistics, see Rizzo (2019); and for a Bayesian perspective, see the "Bayesian Workflow" ideas McElreath (2018).

it your own. That behavior is where deep understanding comes from, not from highlighting the occasional sentence or equation you find important.

Prerequisites

This book is intended for a narrow audience, predominently graduate students with an interest in discrete choice modelling who will find value from seeing and interacting with the programmatic implementation of the multinomial logit and its extended family of related models. In other words, someone who read Train (2009) and thought "how would I code that?"

I assume you are reasonably familiar with R, a freely available software environment for statistical computing and graphics. I also assume you have taken statistics or econometrics courses that introduced the method of maximum likelihood, Bayesian statistics, and limited dependent variable regression models.

Acknowledgments

add:

- R, Posit, Quarto, Sawtooth, and related communities
- Ken, Ella, Peter, Eric, Prachi
- Alison, Zach, Ben

Colophon

An online version of this book is available at https://dcms-r.danyavorsky.com. The source of the book is available at https://github.com/dyavorsky/dcms-r. The book is authored using Quarto, an open-source scientific and technical publishing system that makes it easy to create articles, presentations, websites, book and other publications that combine text and executible code.

Part I

Introduction

1 Introduction

This is a book created from markdown and executable code.

It's the R Companion to the book *Discrete Choice Methods with Simulation* (Train (2009)).

2 Properties of Discrete Choice Models

add

Part II

Behavioral Models

3 Logit

add

4 GEV

add

5 Probit

add

6 Mixed Logit

add

````
`<!-- quarto-file-metadata: eyJyZXNvdXJjZURpciI6ImNoYXB0ZXJzIn0= -->`{=html}

```{=html}
<!-- quarto-file-metadata: eyJyZXNvdXJjZURpciI6ImNoYXB0ZXJzIiwiYm9va0l0ZW1UeXBlIjoiY2hhcHRlc
````
````

7 Variation on a Theme

add

8 Numerical Maximization

add

Part III

Estimation

9 Drawing from Densities

add

10 Simulation-Assisted Estimation

add

11 Individual-Level Parameters

add

12 Bayesian Procedures

add

13 Endogeneity

add

14 EM Algorithms

add

References

McElreath, Richard. 2018. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Chapman; Hall/CRC.

Rizzo, Maria L. 2019. *Statistical Computing with r*. Chapman; Hall/CRC.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press. https://eml.berkeley.edu/books/choice2.html.

A Appendix

add