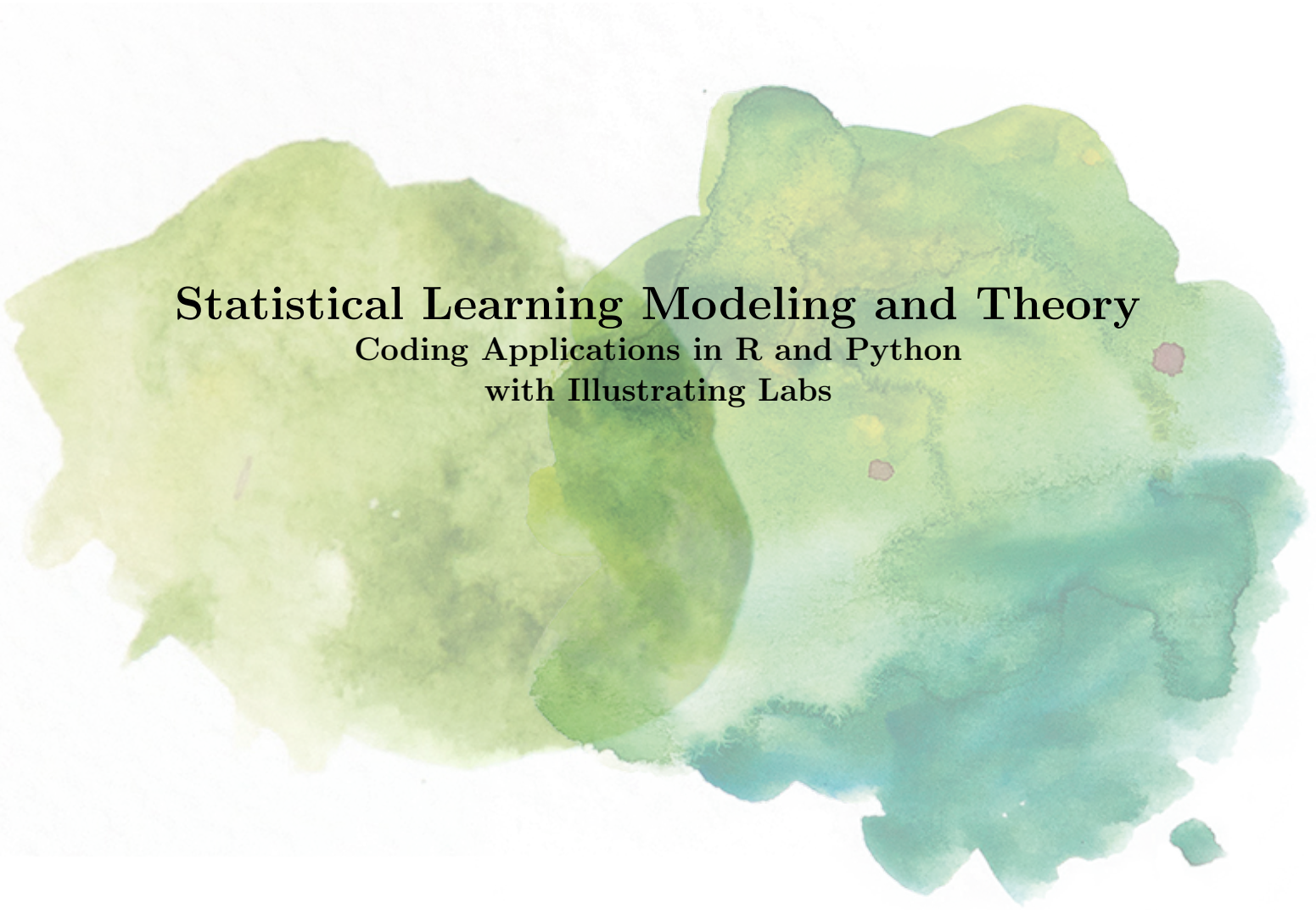

Statistical Learning Modeling and Theory

Coding Applications in R and Python with Illustrating Labs

"Big data isn't about bits,
it's about talent."
– Douglas Merill

Contents

Foreword ii

Acknowledgement ii

1 Introduction 1

2 Statistical Learning 3

2.1 The Basics of Statistical Learning 3

2.1.1 Types of Estimation 3

Chapter 1

Introduction

General Outline for the Intro:

1. The growth of information.
2. The rise of data and AI; "Data is the new oil" - Clive Humby, and "AI is the new electricity" - Andrew Ng
3. How statistical learning fits into the picture; what it does; and why it's important.
4. What the book covers, and what the book can offer.
5. Who the book is "for".
6. How to use the book.

Chapter 2

Statistical Learning

2.1 The Basics of Statistical Learning

Generally speaking, statistical learning is a construct that attempts to predict an outcome, or determine a relationship. Now, in order to make a prediction, or determine an assiciation, you need a minimum two things: an input and output (think: "How does X relate to Y?", or "If I have X, what will be my Y?"). However, more often that not, you may have several things that can influence an output, or a relationship. Hence, we introduce the notation for the input variable(s), X:

$$X = (X_1, X_2, ..., X_P).$$

It is worth taking the time to stop, and make a mental note of the various names that are synonomous to the term: *variable*. Within the scope of statistical learning, the term *variable* is interchangeable with *predictors* (hence the P in X_P), *independent variables*, or *features*. While, the output variable is often referred to as the *response*, or *dependent variable*.

2.1.1 Types of Estimation

Statistical learning can be split into two different kinds of estimation: *prediction* and *inference*. When the time comes, you may determine that one estimation better frames your question than the other. Conversely, you may determine that your question is best answered with the utilization of both types of estimation. Nevertheless, both forms of estimation offer unique insight, and require unique paramaters. We discuss each in depth.

Prediction

Discussion about prediction

Inference

Discussion about inference