Preface

Logistic regression is widely used throughout many scientific disciplines. Your reading of this preface indicates that you are in one of these disciplines. Logistic regression is the prime choice for any research that classifies observations into binary groups, such as success versus failure. To illustrate how widely logistic regression is used, Alan Rowberg suggested that I ask PubMed, so I searched for the term "logistic regression", which produced a list of 52,082 publications at the time of this writing. Although logistic regression is widely used, its results often are misinterpreted or fail to provide the best possible interpretation. In this book, I aim to present the best ways to interpret logistic regression for both professional and lay audiences.

Suppose that you read an article in a medical journal and find a statement that a mother being a smoker significantly increases the likelihood of her newborn having a low birthweight because the coefficient is positive and the p-value is small. The article might list a coefficient and an odds ratio in a table without much interpretation. You should be careful with statements like this. The p-value may be small even when the effect has no practical significance, so it is important to ask questions that may help your interpretation of the results. What is the strength of the impact? What additional risk or reduction in risk does smoking contribute? What graphs would help readers visualize the relationship?

In this book, I will report the p-value of findings, but my focus is on practical significance rather than statistical significance. A small p-value does not mean that the effect has any practical importance, especially with large samples. Because of this, I am following the suggestion of Wasserstein and Lazar (2016) and not using a specific cutoff, such as $p < 0.05$, as evidence of significance.

Do you have enough statistical background to use this book? Yes, if you have

1. completed a single introductory course in statistics and

2. have a basic knowledge of Stata.

Even if you have a more advanced background, this book can still teach you Stata tricks and the best use of logistic regression.

How to use this book. The chapters in this book are short and focus on specific applications of logistic regression. Ideally, you will work through this short book from beginning to end. However, you may wish to refer back to individual chapters as needed, so there is some duplication of content to allow chapters to stand alone. The appendix

includes a do-file that reproduces nearly everything covered in this book. This do-file should be a valuable reference to consult for your research.