

Preface

This is the third edition of this text on logistic regression methods, originally published in 1994, with its second edition published in 2002.

As in the first two editions, each chapter contains a presentation of its topic in “lecture-book” format together with objectives, an outline, key formulae, practice exercises, and a test. The “lecture book” has a sequence of illustrations, formulae, or summary statements in the left column of each page and a script (i.e., text) in the right column. This format allows you to read the script in conjunction with the illustrations and formulae that highlight the main points, formulae, or examples being presented.

This third edition has expanded the second edition by adding three new chapters and a modified computer appendix. We have also expanded our overview of modeling strategy guidelines in Chap. 6 to consider causal diagrams. The three new chapters are as follows:

- Chapter 8: Additional Modeling Strategy Issues
- Chapter 9: Assessing Goodness of Fit for Logistic Regression
- Chapter 10: Assessing Discriminatory Performance of a Binary Logistic Model: ROC Curves

In adding these three chapters, we have moved Chaps. 8 through 13 from the second edition to follow the new chapters, so that these previous chapters have been renumbered as Chaps. 11–16 in this third edition. To clarify this further, we list below the previous chapter titles and their corresponding numbers in the second and third editions:

<i>Chapter Title</i>	<i>Chapter # 2nd Edition</i>	<i>Chapter # 3rd Edition</i>
Analysis of Matched Data Using Logistic Regression	8	11
Polytomous Logistic Regression	9	12
Ordinal Logistic Regression	10	13
Logistic Regression for Correlated Data: GEE	11	14
GEE Examples	12	15
Other Approaches for Analysis of Correlated Data	13	16

New Chap. 8 addresses five issues on modeling strategy not covered in the previous two chapters (6 and 7) on this topic:

Issue 1: Modeling Strategy When There Are Two or More Exposure Variables

Issue 2: Screening Variables When Modeling

Issue 3: Collinearity Diagnostics

Issue 4: Multiple Testing

Issue 5: Influential Observations

New Chap. 9 addresses methods for assessing the extent to which a binary logistic model estimated from a dataset predicts the observed outcomes in the dataset, with particular focus on the deviance statistic and the Hosmer-Lemeshow statistic.

New Chap. 10 addresses methods for assessing the extent that a fitted binary logistic model can be used to distinguish the observed cases from the observed noncases, with particular focus on ROC curves.

The modified appendix, Computer Programs for Logistic Regression, updates the corresponding appendix from the second edition. This appendix provides computer code and examples of computer programs for the different types of logistic models described in this third edition. The appendix is intended to describe the similarities and differences among some of the most widely used computer packages. The software packages considered are SAS version 9.2, SPSS version 16.0, and Stata version 10.0

Suggestions for Use

This text was originally intended for self-study, but in the 16 years since the first edition was published, it has also been effectively used as a text in a standard lecture-type classroom format. Alternatively, the text may be used to supplement material covered in a course or to review previously learned material in a self-instructional or distance-learning format. A more individualized learning program may be particularly suitable to a working professional who does not have the time to participate in a regularly scheduled course.

The order of the chapters represents what the authors consider to be the logical order for learning about logistic regression. However, persons with some knowledge of the subject can choose whichever chapter appears appropriate to their learning needs in whatever sequence desired.

The last three chapters (now 14–16) on methods for analyzing correlated data are somewhat more mathematically challenging than the earlier chapters, but have been written

to logically follow the preceding material and to highlight the principal features of the methods described rather than to give a detailed mathematical formulation.

In working with any chapter, the user is encouraged first to read the abbreviated outline and the objectives, and then work through the presentation. After finishing the presentation, the user is encouraged to read the detailed outline for a summary of the presentation, review key formulae and other important information, work through the practice exercises, and, finally, complete the test to check what has been learned.

Recommended Preparation

The ideal preparation for this text is a course on quantitative methods in epidemiology and a course in applied multiple regression. The following are recommended references on these subjects with suggested chapter readings:

Kleinbaum, D.G., Kupper, L.L., and Morgenstern, H., *Epidemiologic Research: Principles and Quantitative Methods*, Wiley, New York, 1982, Chaps. 1–19.

Kleinbaum, D.G., Kupper, L.L., Nizam, A., and Muller, K.A., *Applied Regression Analysis and Other Multivariable Methods, Fourth Edition*, Duxbury Press/Cengage Learning, Pacific Grove, 2008, Chaps. 1–16.

Kleinbaum, D.G., *ActivEpi- A CD-Rom Text*, Springer, New York, 2003, Chaps. 3–15.

A first course on the principles of epidemiologic research would be helpful since this text is written from the perspective of epidemiologic research. In particular, the learner should be familiar with basic characteristics of epidemiologic study designs and should have some understanding of the frequently encountered problem of controlling/adjusting for variables.

As for mathematics prerequisites, the learner should be familiar with natural logarithms and their relationship to exponentials (powers of e) and, more generally, should be able to read mathematical notation and formulae.

Atlanta, GA

David G. Kleinbaum
Mitchel Klein



<http://www.springer.com/978-1-4419-1741-6>

Logistic Regression

A Self-Learning Text

Kleinbaum, D.G.; Klein, M.

2010, XVII, 702 p., Hardcover

ISBN: 978-1-4419-1741-6