Kenneth Knoblauch, Laurence T. Maloney

Modeling Psychophysical Data in R

– Monograph –

April 12, 2012

Springer

Preface

This book is about modeling psychophysical data with modern statistical methods using the software language R [146]. R is an interactive environment and the novice user can just enter commands one by one at the command line prompt. It is also a powerful programming language in its own right. R is easily interfaced with other languages, such as C, Fortran, C++, Java, Perl and Python. Another reason we recommend using R is that it permits access to a growing repository of new and powerful statistical methods many of which are potentially useful to the psychophysical researcher. In the last chapter, for example, we describe methods that allow the researcher to model individual differences in observers, using mixed-effect's models.

Many packages exist to supplement the functions of R. Over 3000 are currently available on the Comprehensive R Archive Network (or CRAN) at `http://www.r-project.org/`. Further reasons for using R are that it is free and is supported by a large and enthusiastic online community.

Psychophysics begins with a question, posed to a *psychophysical observer*, human, animal, or neural[1]. The question may be as simple as judging which of two lights is brighter. The experimenter selects the question, constrains the observer's possible responses, and manipulates stimuli. The observer's response is effectively stochastic and the experimenter's proximate goal is to model the observer's response and the mechanisms underlying that response. The final goal, of course, is to learn how neural systems process sensory information and provide access to the world around us.

The term "psychophysics" is due to Fechner [59], and there are several excellent introductions to psychophysical methods [91, 157, 158, 159, 164], the history of psychophysics [17, 115], and psychophysical theory [57, 151]. We note that psychophysics has no connection with psychometric theory. For an excellent introduction to psychometric theory using R we recommend Boeck & Wilson [16].

[1] We can treat any device that categorizes stimuli as a psychophysical observer. See [22, 138, 174] for examples of psychophysical models of neural response.

This book is *not* an introduction to psychophysical methods, history or theory. It is written for those who have a good idea what psychophysics entails and who regularly model psychophysical data or would like to do so—typically established or aspiring psychologists and neuroscientists studying perception. Some of the material is pertinent to researchers in memory using signal detection theory.

The focus of the book is on modern statistical methods readily available in R and how they can be used to analyze psychophysical data. Throughout the book, we illustrate how traditional and recently developed models of psychophysical performance can be viewed as special cases of statistical models of great power and generality, notably the Generalized Linear Model (GLM). In Chap. 1, we use R to explore and model a celebrated set of psychophysical data due to Hecht, Shlaer & Pirenne [79]. When possible, as in this chapter, we use a running example within the chapter so that readers can appreciate how different methods might be applied to the same data set. Thus, packages and data sets loaded earlier in the chapter as well as variables will be assumed to be defined from the point in the chapter in which they are invoked, but not across chapters.

The statistical methods employed in most psychophysical laboratories were developed for the most part over the past 50 years and until recently, commonly available statistical packages were of limited use in fitting psychophysical data. That has now changed and much of what the psychophysicist wishes to learn from data and experiment can be learned using off-the-shelf statistical packages developed in the last few decades. Moreover, the statistical methods and packages now available can serve to inspire the psychophysicist to try more sophisticated analyses and consider better experimental designs. If, for example, the experimenter has run an experiment with a factorial design then he might like to do the non-linear equivalent of an analysis of variance (ANOVA) and determine whether variation of any of the factors led to changes in observers' performance and whether variations in the factors interact. We illustrate how to carry out ANOVA-like analyses where the data are appropriate for fitting by signal detection models or psychometric functions.

The book should be accessible to upper level undergraduate and graduate students and could be used as a textbook for a one-semester course. It presupposes only some knowledge of elementary statistics, some facility with vector and matrix notation, and some familiarity with programming. We do not assume that the reader is familiar with R or is particularly fluent in any other programming language. Almost all of the examples are drawn from studies of human vision only because that is what we know best. The methods presented are readily adaptable to model data in other disciplines ranging from memory to experimental economics and in testing this book with students we have found that they could readily adapt the code we include as examples to their own data without any extensive programming.

Chapter 2 describes the principal statistical methods we will use: the linear model (LM), the Generalized Linear model (GLM)[2] the (Generalized) Additive Model (GAM) and a brief discussion of methods for fitting nonlinear models in general. The reader familiar with LM in the form of multiple regression or analysis of variance can become familiar with how models are specified and analyzed in R. The remainder of this chapter introduces GLM and GAM by analogy with LM. The treatment of GLM and GAM in this book is self-contained. We provide references to texts on LM, GLM and GAM using R for readers who want to pursue these topics further. If already familiar with R and its approach to modeling, the reader may choose to skip this chapter on an initial reading and to start instead with one of the later chapters describing particular psychophysical methods. The reader can refer back to Chap. 2 as needed.

We also include two appendices that summarize a few useful terms and results from mathematical statistics (Appendix B) and a brief description of data structures in R as well as basic information about input and output of data in R (Appendix A). We reference particular sections of these appendices in the main text, and they need not be read through if one is already familiar with their content.

The remaining chapters each focus on traditional and not so traditional methods of fitting data to psychophysical models using modern statistical methods. With the exception of Chap. 5 which builds on Chap. 4, the chapters are independent of one another and may be read in any order. We recommend that, wherever you start, you type in and execute the examples in the text as you read them. All of the code is available in scripts in a companion R package **MPDiR** available on CRAN that also contains the data sets not available elsewhere. There are exercises at the end of each chapter that invite the reader to modify the example code or to explore further the R commands used. If you are motivated to go further, you can apply the example code to your own data. Even without extensive experience in programming or programming in R, you should be able to adapt your own data to our examples. All of the data analyses and graphs in this book were generated using R and we encourage the reader to execute the examples in the book as they occur. In our experience, using R is the best way to learn R.

In Chap. 3 we describe how to use GLM to fit data from signal detection experiments. Signal Detection Theory (SDT) was introduced in the 1950"s to address the problem of observer bias. In a simple Yes-No task, a psychometric function summarizes the probability of a 'Yes' response. However, this probability confounds the observer's ability to detect the presence or absence of a signal with any preference the observer may have to respond 'Yes'. Even when the signal intensity is zero (no signal is presented), the observer may still respond 'Yes'. In a signal detection experiment, in simplest form, there is

[2] The Generalized Linear Model is *not* the same as the General Linear Model commonly used in cognitive neuroscience. The near-coincidence in terminology is unfortunate. We will not make use of the General Linear Model at any point.

only one intensity of signal that is presented or not presented, at random, on each trial. On each trial, the observer is obliged to judge whether a signal was present. The two measures of interest are the rate of responding "Present" when the signal is present (referred to as the Hit rate) and the rate of responding "Present" when the signal is absent, the False Alarm rate. These two separate response measures are combined to allow separate estimates of a measure of the observer's sensitivity to the signal and the observer"s bias (tendency to say "Present" independent of the presence or absence of the signal).

In Chap. 4 we illustrate how to fit psychometric functions to data using GLM. The psychometric function describes the relation between the probability of making a specific classification such as "Present" and stimulus intensity. Its earliest formulation dates from at least the beginning of the twentieth century [176]. One long-established use of psychophysical methods is to measure the precision of the human sensory apparatus in detection or discrimination. The researcher's goal nowadays is often to use such data to test models of sensory, perceptual or cognitive function and one of the most venerable is the *psychometric function* relating the observer's ability to judge a stimulus to stimulus intensity. There are many variant forms of psychometric function each associated with a particular experimental method. For now we assume that the observer responds "Yes" or "No", judging whether the stimulus was present.

In Chap. 5 we extend the analysis of the psychometric function to more complex situations, for example, when there are multiple choices. We, also, consider here methods for obtaining standard errors and confidence intervals on the parameters estimated for the fitted functions. Finally, we demonstrate some nonparametric approaches to estimating psychometric functions.

In Chap. 6 we describe classification image methods. These methods are a natural extension of signal detection theory and have proven to be of great value in modeling psychophysical data. They are also easy to work with using GLM and GAM.

In the next two chapters, we describe two relatively new methods for representing perception of stimulus intensities well above threshold. That is, these methods are primarily useful in measuring *appearance*. The first method, Maximum Likelihood Difference Scaling, is described in Chap. 7, the second, Maximum Likelihood Conjoint Measurement, in Chap. 8. These methods and corresponding psychophysical tasks complement traditional methods. We will illustrate their use.

In many statistical analyses the experimenter wishes to estimate how variable the outcomes of psychophysical measurements are across observers, generalized from results measured for a specific set of observers in an experiment to the population from which they were drawn. In the last chapter, Chap. 9 we focus on how to use newly-available R packages to fit mixed-effects models of psychophysical performance. These models take us to the edge of research in statistics and hint at what is to come.

The more advanced programmer will want to know that R is an object-oriented language. Most of the functions that carry out statistical analyses return not just vectors or matrices, but objects summarizing the analyses carried out. Generic functions that print, plot, summarize or do model comparisons accept these objects and, choose the appropriate methods to do what is appropriate. In general, we simply use these functions as needed to illustrate how intuitive they are.

R is a language that is in constant development and although the core functionality evolves only slowly over time, usually with an eye toward backward compatibility, this is less so for the code in packages. In time, some of the code in this book will be affected by such changes and may not run. We have opted to choose package versions that seem stable but in the event that future changes break some of the code, we will make available supplemental materials on a web page associated with the book that can be accessed at `http://extras.springer.com/`. The current project home page for the book can be found at `http://mpdir.r-forge.r-project.org/`.

We thank Angela Brown, Frédéric Devinck, Roland Fleming, Jonathan Pierce, Hao Sun, François Vital-Durand and Li Zhaoping for generously providing data. We thank Rune H. B. Christensen for extended discussions of ordinal models and Aries Arditi, Patrick Bennett, Reinhold Kleigl, Richard Murray, and, above all, Michael Kubovy for reading and critiquing several chapters. We thank our contacts at Springer, the editors John Kimmel and Marc Strauss, and Hannah Bracken. We also thank all of the R Core and R Help List contributors who responded to many questions along the way. Any remaining errors are, of course, our own.

We especially thank Geneviève Inchauspé and Maria Felicita Dal Martello for infinite patience with our preoccupation with R and this book.

Lyon, New York and Padova — *Kenneth Knoblauch*
April 12, 2012 — *Laurence T. Maloney*

Contents

1 A First Tour through R by Example 1
- 1.1 Getting started 1
- 1.2 The experiment of Hecht, Shlaer & Pirenne 1
  - 1.2.1 Accessing the data set 2
  - 1.2.2 Viewing the data set 5
  - 1.2.3 One more plot before modeling 7
- 1.3 Modeling the data 8
  - 1.3.1 Modeling a subset of the data 10
  - 1.3.2 Estimating thresholds and jnd's 15
  - 1.3.3 Modeling all of the data 17
  - 1.3.4 Odds and ends 18
- 1.4 Exercises 20

2 Modeling in R 21
- 2.1 Introduction 21
- 2.2 The Linear Model 22
  - 2.2.1 Development of first and second order motion 26
- 2.3 Linear mixed-effects models 32
  - 2.3.1 The ModelFest data set 34
- 2.4 Nonlinear models 43
  - 2.4.1 Chromatic sensitivity across the life span 45
- 2.5 Additive models 50
  - 2.5.1 Chromatic sensitivity, again 51
- 2.6 Generalized linear models 54
  - 2.6.1 Chromatic sensitivity, yet again 55
- 2.7 Which model to choose? 57
- 2.8 Exercises 58

3 Signal Detection Theory (SDT) 61
- 3.1 Introduction 61
- 3.2 Sensory and decision spaces 61

3.2.1 Optimal detection . . . 63
3.3 Case 1: Equal-variance Gaussian SDT . . . 64
3.3.1 The decision rule for equal-variance Gaussian SDT . . . 66
3.3.2 Simulating and fitting equal variance Gaussian SDT . 66
3.3.3 Maximum likelihood estimation of equal variance Gaussian SDT parameters . . . 68
3.3.4 Fisher's tea test . . . 69
3.3.5 Fitting equal-variance SDT as a GLM . . . 73
3.3.6 Fitting with multiple conditions . . . 75
3.3.7 Complete separation . . . 78
3.3.8 The Receiver Operating Characteristic . . . 79
3.4 Case 2: Unequal-variance Gaussian SDT . . . 81
3.4.1 The decision rule for unequal-variance Gaussian SDT . 82
3.4.2 Simulating unequal-variance Gaussian SDT . . . 82
3.4.3 Fitting unequal-variance Gaussian SDT by direct optimization . . . 86
3.5 Case 3: Exponential SDT . . . 88
3.6 m-Alternative Forced-Choice (mAFC) and other paradigms . . 90
3.6.1 Letter identification: 4AFC . . . 91
3.6.2 d' for other paradigms . . . 92
3.7 Rating scales . . . 92
3.7.1 Ordinal regression models on cumulative probabilities . 96
3.7.2 Sensitivity to family resemblance . . . 100

4 The Psychometric Function: Introduction . . . 107
4.1 Brief historical review . . . 107
4.2 Fitting psychometric functions to data . . . 113
4.2.1 Direct maximization of likelihood . . . 114
4.3 A model psychophysical observer . . . 120
4.4 Fitting psychometric functions with `glm` . . . 122
4.4.1 Alignment judgments of drifting grating pairs . . . 123
4.4.2 Selecting models and analysis of deviance . . . 126
4.5 Diagnostic plots . . . 130
4.6 Complete separation . . . 136
4.7 The fitted model . . . 137
4.8 Summary . . . 138

5 The Psychometric Function: Continuation . . . 143
5.1 Fitting Multiple-Alternative Forced Choice (mAFC) data sets 143
5.1.1 Using an offset to set the lower asymptote . . . 144
5.1.2 Using specialized link functions to set the lower asymptote . . . 147
5.1.3 Estimating upper and lower asymptotes . . . 149
5.2 Staircase methods . . . 153
5.3 Extracting fitted parameters . . . 155

5.4 Methods for standard errors and confidence intervals 159
5.5 Bootstrap standard error estimates 161
5.6 Nonparametric approaches 167

6 Classification Images 173
6.1 Ahumada's experiment 173
6.1.1 Simulating a classification image experiment 175
6.1.2 The information used in offset detection 179
6.2 Some history 180
6.3 The classification image as a linear model 181
6.3.1 Detection of a Gabor temporal modulation 182
6.3.2 Estimation by a linear model 183
6.3.3 Estimation with `lm` 184
6.3.4 Model evaluation 185
6.4 Classification image estimation with `glm` 186
6.5 Fitting a smooth model 190
6.5.1 Higher-order classification images 194
6.6 Comparing methods 196

7 Maximum Likelihood Difference Scaling (MLDS) 203
7.1 Introduction 203
7.1.1 Choosing the quadruples 206
7.1.2 Data format 208
7.1.3 Fitting the model: Direct maximization 209
7.1.4 GLM method 212
7.1.5 Representation as a ratio scale 214
7.2 Perception of correlation in scatterplots 214
7.2.1 Estimating a difference scale 216
7.2.2 Fitting a parametric difference scale 218
7.2.3 Visualizing the decision space for difference judgments 219
7.3 The Method of Triads 223
7.4 The effect of number of trials 224
7.5 Bootstrap standard errors 227
7.6 Robustness and diagnostic tests 228
7.6.1 Goodness of fit 229
7.6.2 Diagnostic test of the measurement model 232

8 Maximum Likelihood Conjoint Measurement (MLCM) . . 239
8.1 Introduction 239
8.2 The model 243
8.3 The nonparametric model 247
8.3.1 Fitting the nonparametric model directly by maximum likelihood 251
8.3.2 Fitting the nonparametric model with `glm` 253
8.4 Fitting parametric models 258

8.5 Resampling methods . . . 263
8.5.1 The choice of stimuli . . . 264
8.6 The axiomatic model . . . 265

9 Mixed-Effect Models . . . 269
9.1 Simple Detection . . . 271
9.1.1 Random criterion effects . . . 271
9.1.2 Random sensitivity effects . . . 275
9.2 Rating scale data . . . 276
9.3 Psychometric functions: Random intercepts and slopes . . . 280
9.3.1 Context effects on detection . . . 289
9.3.2 Multiple-alternative forced-choice data . . . 298
9.4 MLDS . . . 299
9.4.1 Regression on the decision variable: Fixed functional form . . . 300
9.4.2 Regression on the decision variable: Rescaling to a common scale . . . 303
9.4.3 Regression on the estimated coefficients . . . 308
9.5 Whither from here . . . 314

A Some Basics of R . . . 315
A.1 Installing R . . . 315
A.2 Basic Data Types . . . 316
A.2.1 Vectors . . . 316
A.2.2 Factors . . . 322
A.2.3 Lists . . . 323
A.2.4 Data frames . . . 324
A.2.5 Other types of objects . . . 326
A.3 Simple programming constructs . . . 326
A.4 Interacting with the operating system . . . 327
A.5 Getting data into and out of R . . . 327
A.5.1 Writing and reading text files . . . 328
A.5.2 `write.table` . . . 328
A.5.3 `read.table` . . . 328
A.5.4 `scan` . . . 330
A.5.5 `save` and `load` . . . 330
A.5.6 Interacting with Matlab . . . 331
A.5.7 Interacting with Excel files . . . 331
A.5.8 Reading in other formats . . . 332

B Statistical Background . . . 333
B.1 Introduction . . . 333
B.2 Random variables and Monte Carlo methods . . . 335
B.2.1 Finite, discrete random variables . . . 336
B.2.2 Bernoulli random variables . . . 336

B.2.3 Continuous random variables 336
B.2.4 Discrete vs. continuous random variables 338
B.2.5 Normal (Gaussian) random variables 338
B.2.6 Location-scale families 339
B.3 Independence and samples 340
B.3.1 Expectation and moments 342
B.3.2 Bernoulli and binomial random variables 343
B.3.3 Uniform random variables 344
B.3.4 Exponential and Gamma random variables 345
B.3.5 The exponential family 347
B.3.6 Monte Carlo methods 348
B.4 Estimation 348
B.4.1 What makes an estimate "good"? 349
B.4.2 Maximum likelihood estimation 350
B.4.3 Why likelihood? Some properties 353
B.4.4 Bayesian estimation 355
B.5 Nested hypothesis testing 361
B.5.1 Reparameterization: Nested hypothesis tests 363
B.5.2 The Akaike Information Criterion (AIC) 364

References 365

Index 375