

Statistics for Social and Behavioral Sciences

Series editor

Stephen E. Fienberg
Carnegie Mellon University
Pittsburgh
Pennsylvania
USA

Statistics for Social and Behavioral Sciences (SSBS) includes monographs and advanced textbooks relating to education, psychology, sociology, political science, public policy, and law.

More information about this series at <http://www.springer.com/series/3463>

Frank B. Baker • Seock-Ho Kim

The Basics of Item Response Theory Using R

 Springer

Frank B. Baker
Educational Psychology
University of Wisconsin-Madison
Madison, WI, USA

Seock-Ho Kim
Educational Psychology
University of Georgia
Athens, GA, USA

ISSN 2199-7357 ISSN 2199-7365 (electronic)
Statistics for Social and Behavioral Sciences
ISBN 978-3-319-54204-1 ISBN 978-3-319-54205-8 (eBook)
DOI 10.1007/978-3-319-54205-8

Library of Congress Control Number: 2017933685

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book is an update of the original book, *The Basics of Item Response Theory*, by the first author. The original book by Frank B. Baker was based on the course given during his tenure at the University of Wisconsin–Madison. It appeared in 1985. The second edition of the original book by Frank B. Baker appeared in 2001 with a publisher’s note by Lawrence A. Rudner. About 15 years have passed since the last revision. So much has happened in the fields of educational measurement and psychometrics as well as in the statistical computing technology. In the meantime, we felt that the usefulness of the book would be increased by some further changes. The main alterations are due to the use of the computing package R for the illustration purpose and especially for the computer sessions. The treatment of the original topics over eight chapters has not been changed.

The original object of the book was to make the book to be a tutorial for item response theory suited to those who possess only a limited knowledge of educational measurement and psychometrics. We have never lost sight of such an object. The amendments in this book are not due to any alteration in the original object but they are necessitated by the development of the statistical computing technology. In particular, the book now aims at covering both the basics of item response theory and the use of R for preparing graphical presentation in the item response theory related writings.

We will be indebted to any reader who calls our attention to errors or obscurities.

Madison, WI, USA
Athens, GA, USA
January 2017

Frank B. Baker
Seock-Ho Kim

Acknowledgments of the Original Book (1985)

Over the past century, many people have contributed to the development of item response theory. Three persons deserve special recognition. D. N. Lawley of the University of Edinburgh published a paper in 1943 showing that many of the constructs of classical test theory could be expressed in terms of parameters of the item characteristic curve. This paper marks the beginning of item response theory as a measurement theory. Dr. F. M. Lord of the Educational Testing Service has been the driving force behind both the development of theory and its application for the past 35 years. Over this period, he has systematically defined, expanded, and explored the theory as well as developed the computer programs needed to put the theory into practice. This effort culminated in his recent book on the practical applications of item response theory. In the late 1960s, Dr. B. D. Wright of the University of Chicago recognized the importance of the measurement work by the Danish mathematician Georg Rasch. Since that time, he has played a key role in bringing item response theory, the Rasch model in particular, to the attention of practitioners. Without the work of these three individuals, the level of development of item response theory would not be where it is today.

I am indebted to Mr. T. Seavey of Heinemann Educational Books for first suggesting that I do a small book on item response theory. This suggestion allowed me to fulfill a long-standing desire to develop an instructional software package dealing with item response theory for a microcomputer. I must also acknowledge the technical assistance of Mr. W. Vilberg in squeezing the maximum capability out of the Apple II computer. Without his help, the computer software would be much less sophisticated. Finally, the manuscript was prepared using the Screenwriter II word processor program written by Mr. R. Kidwell of Sierra On-Line, Inc. Without this marvelous package, the present book would never be written.

Madison, WI, USA

Frank B. Baker

Acknowledgements of the Second Edition of the Original Book (2001)

Over the past century, many people have contributed to the development of item response theory. Three persons deserve special recognition. D. N. Lawley of the University of Edinburgh published a paper in 1943 showing that many of the constructs of classical test theory could be expressed in terms of parameters of the item characteristic curve. This paper marks the beginning of item response theory as a measurement theory. The work of Dr. F. M. Lord of the Educational Testing Service has been the driving force behind both the development of theory and its application for the past 50 years. Dr. Lord systematically defined, expanded, and explored the theory as well as developed the computer programs needed to put the theory into practice. This effort culminated in his classic books (with Dr. Melvin Novick, 1968; 1980) on the practical applications of item response theory. In the late 1960s, Dr. B. D. Wright of the University of Chicago recognized the importance of the measurement work by the Danish mathematician Georg Rasch. Since that time, he has played a key role in bringing item response theory, the Rasch model in particular, to the attention of practitioners. Without the work of these three individuals, the level of development of item response theory would not be where it is today.

I am indebted to Mr. T. Seavey of Heinemann Educational Books for first suggesting that I do a small book on item response theory, which resulted in the first edition of this book in 1985. This suggestion allowed me to fulfill a long-standing desire to develop an instructional software package dealing with item response theory for the then-state-of-the-art Apple II and IBM PC computers. An upgraded version of this software has now been made available on the World Wide Web (<http://wricae.net/irt>).

Madison, WI, USA

Frank B. Baker

Contents

1	The Item Characteristic Curve	1
2	Item Characteristic Curve Models	17
3	Estimating Item Parameters	35
4	The Test Characteristic Curve	55
5	Estimating an Examinee’s Ability	69
6	The Information Function	89
7	Test Calibration	105
8	Specifying the Characteristics of a Test	127
A	R Introduction	137
B	Estimating Item Parameters Under the Two-Parameter Model with Logistic Regression	165
C	Putting the Three Tests on a Common Ability Scale: Test Equating ...	167
	References	171
	Index	173

Introduction

When the original book was first published in 1985, the fields of educational measurement and psychometrics were in a transitional period. The majority of practice was based upon the classical test theory developed during the 1920s. However, a new test theory had been developing over the past 40 years that was conceptually more powerful than classical test theory. Based upon items rather than test scores, the new approach was known as item response theory. While the basic concepts of item response theory are straightforward, the underlying mathematics is somewhat advanced compared to that of classical test theory. As a result, it is difficult to examine some of these concepts without performing a large number of calculations to obtain usable information. The original book was designed to provide the reader access to the basic concepts of item response theory freed of the tedious underlying calculations through an Apple II computer program. The second edition of the original published in 2001 used a version of computer program written in Visual Basic 5.0 that could be obtained at <http://ericae.net/irt>. Readers accustomed to sophisticated statistical and graphics packages might have found it utilitarian, but nevertheless helpful in understanding various facets of the theory. This book now uses R that is a freely available programming language for applied statistics and data visualization. The file folder accompanying the book contains a set of R functions that implement various facets of the theory. These R functions allow the reader to explore the theory at the conceptual level.

The book is organized in a building block fashion. It proceeds from the simple to the complex with each new topic building on the preceding topics. Within each of the eight chapters, a basic concept is presented, the corresponding computer session is explained, and a set of exploratory exercises are defined. Readers are then strongly encouraged to use the computer session to explore the concept through a series of exercises. A final section of each chapter, called “Things to Notice,” lists some of the characteristics of the concept that you should have noticed and some of the conclusions that you should have reached. If you do not understand the logic underlying something in this section, you can return to the computer session and try new variations and explorations until clarity is achieved.

When finished with the book and the computer sessions, the reader should have a good working knowledge of the fundamentals of item response theory. This book emphasizes the basics, minimizes the amount of mathematics, and does not pursue technical details that are of interest only to the specialist. In some sense, you will be shown only “what you need to know” rather than all the glorious details of the theory. Upon completion of this book, the reader should be able to interpret test results that have been analyzed under item response theory by means of programs such as WINSTEPS (Linacre 2015), BILOG-MG (Zimowski et al. 2002), and PCLOGIST (Wingersky et al. 1999). Note that WINSTEPS is a current descendant of BICAL (Wright and Mead 1976), BILOG-MG is the extended version of BILOG (Mislevy and Bock 1984), and PCLOGIST is the personal computer version of LOGIST (Wingersky et al. 1982). In order to employ the theory in a practical setting, the reader should study more advanced books on the applications of the theory such as Baker and Kim (2004), de Ayala (2009), Embretson and Reise (2000), Nering and Ostini (2010), Reckase (2009), Thissen and Wainer (2001), and van der Linden and Glas (2000) as well as some earlier books including Hambleton and Swaminathan (1984), Hambleton et al. (1991), Wright and Stone (1979), and Hulin et al. (1983).

Getting Started

R is a software package for data analysis and graphical representation. R provides the language, functions, and the computing environment in one convenient package. The main uniform resource locator (URL; i.e., webpage) of R is

<http://www.r-project.org>

You can download R by clicking one of the Comprehensive R Archive Network (CRAN) mirror sites in

<http://cran.r-project.org/mirrors.html>

and following the instruction shown on your computer screen for your own operating system of Linux, Macintosh, or Windows. The `base` subdirectory contains the binaries for R. Appendix A contains a brief introductory summary of R.

After installing R on your computer, you can perform all activities shown in the respective chapters and the computer sessions by typing in the R command lines exactly shown in the book. Alternatively, for larger R command lines that may contain R functions in the book, you can obtain and use a zipped file folder (BIRTRFunctions.zip) that contains R scripts from the publisher’s web site.