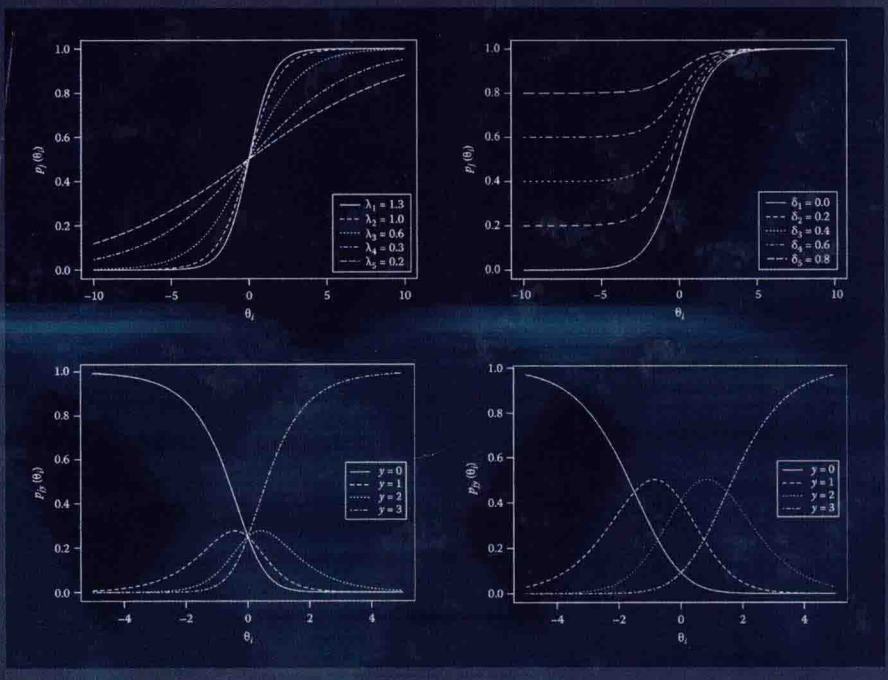
Chapman & Hall/CRC
Interdisciplinary Statistics Series

Statistical Analysis of Questionnaires

A Unified Approach Based on R and Stata



Francesco Bartolucci
Silvia Bacci
Michela Gnaldi

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Chapman & Hall/CRC
Interdisciplinary Statistics Series

Statistical Analysis of Questionnaires

A Unified Approach Based on R and Stata

Francesco Bartolucci
Silvia Bacci
Michela Gnaldi

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper by CPI Group (UK) Ltd, Croydon
Version Date: 20150528

International Standard Book Number-13: 978-1-4665-6849-5 (Hardback)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Statistical Analysis of Questionnaires

A Unified Approach Based on R and Stata

CHAPMAN & HALL/CRC

Interdisciplinary Statistics Series

Series editors: N. Keiding, B.J.T. Morgan, C.K. Wikle, and P. van der Heijden

Published titles

AGE-PERIOD-COHORT ANALYSIS: NEW MODELS, METHODS, AND EMPIRICAL APPLICATIONS Y. Yang and K.C. Land

ANALYSIS OF CAPTURE-RECAPTURE DATA R.S. McCrea and B.J.T. Morgan

AN INVARIANT APPROACH TO STATISTICAL ANALYSIS OF SHAPES S. Lele and J. Richtsmeier

ASTROSTATISTICS G. Babu and E. Feigelson

BAYESIAN ANALYSIS FOR POPULATION ECOLOGY R. King, B.J.T. Morgan, O. Gimenez, and S.P. Brooks

BAYESIAN DISEASE MAPPING: HIERARCHICAL MODELING IN SPATIAL EPIDEMIOLOGY, SECOND EDITION A.B. Lawson

BIOEQUIVALENCE AND STATISTICS IN CLINICAL PHARMACOLOGY S. Patterson and B. Jones

CLINICAL TRIALS IN ONCOLOGY, THIRD EDITION S. Green, J. Benedetti, A. Smith, and J. Crowley

CLUSTER RANDOMISED TRIALS R.J. Hayes and L.H. Moulton

CORRESPONDENCE ANALYSIS IN PRACTICE, SECOND EDITION M. Greenacre

DESIGN AND ANALYSIS OF QUALITY OF LIFE STUDIES IN CLINICAL TRIALS, SECOND EDITION D.L. Fairclough

DYNAMICAL SEARCH L. Pronzato, H.Wynn, and A. Zhigljavsky

FLEXIBLE IMPUTATION OF MISSING DATA S. van Buuren

GENERALIZED LATENT VARIABLE MODELING: MULTILEVEL, LONGITUDINAL, AND STRUCTURAL EQUATION MODELS A. Skrondal and S. Rabe-Hesketh

GRAPHICAL ANALYSIS OF MULTI-RESPONSE DATA K. Basford and J. Tukey

INTRODUCTION TO COMPUTATIONAL BIOLOGY: MAPS, SEQUENCES, AND GENOMES M. Waterman

MARKOV CHAIN MONTE CARLO IN PRACTICE W. Gilks, S. Richardson, and D. Spiegelhalter

MEASUREMENT ERROR AND MISCLASSIFICATION IN STATISTICS AND EPIDEMIOLOGY: IMPACTS AND BAYESIAN ADJUSTMENTS P. Gustafson

MEASUREMENT ERROR: MODELS, METHODS, AND APPLICATIONS J.P. Buonaccorsi

MENDELIAN RANDOMIZATION: METHODS FOR USING GENETIC VARIANTS IN CAUSAL ESTIMATION S. Burgess and S.G. Thompson

Published titles

META-ANALYSIS OF BINARY DATA USINGPROFILE LIKELIHOOD D. Böhning, R. Kuhnert, and S. Rattanasiri

POWER ANALYSIS OF TRIALS WITH MULTILEVEL DATA M. Moerbeek and S. Teerenstra

STATISTICAL ANALYSIS OF GENE EXPRESSION MICROARRAY DATA T. Speed

STATISTICAL ANALYSIS OF QUESTIONNAIRES: A UNIFIED APPROACH BASED ON R AND STATA F. Bartolucci, S. Bacci, and M. Gnaldi

STATISTICAL AND COMPUTATIONAL PHARMACOGENOMICS R. Wu and M. Lin

STATISTICS IN MUSICOLOGY J. Beran

STATISTICS OF MEDICAL IMAGING T. Lei

STATISTICAL CONCEPTS AND APPLICATIONS IN CLINICAL MEDICINE J. Aitchison, J.W. Kay, and I.J. Lauder

STATISTICAL AND PROBABILISTIC METHODS IN ACTUARIAL SCIENCE P.J. Boland

STATISTICAL DETECTION AND SURVEILLANCE OF GEOGRAPHIC CLUSTERS P. Rogerson and I. Yamada

STATISTICS FOR ENVIRONMENTAL BIOLOGY AND TOXICOLOGY A. Bailer and W. Piegorsch

STATISTICS FOR FISSION TRACK ANALYSIS R.F. Galbraith

VISUALIZING DATA PATTERNS WITH MICROMAPS D.B. Carr and L.W. Pickle

Preface

Questionnaires are fundamental tools of data collection in many fields and, in particular, in education, psychology, and sociology. They are becoming important tools of data collection even in economics, management, and medicine, especially when an individual characteristic that is not directly observable, such as the level of satisfaction of a customer or the quality of life of a patient, is of main interest.

A questionnaire is typically composed of a series of items that are scored on a binary or ordinal scale and are obviously correlated within each respondent. Therefore, data collected by a questionnaire require special statistical methods to be analyzed. This area has been the object of an increasing interest in the past decades, soliciting a growing number of academic research studies focused on theoretical and practical advances. In particular, in this book, we deal with the foundations of classical test theory (CTT), item response theory (IRT) fundamentals, IRT for dichotomous and polytomous items, up to the most recent IRT extensions, such as multidimensional IRT models, and latent class IRT models. Estimation methods and diagnostics—including graphical diagnostic tools, parametric and nonparametric tests, and differential item functioning—are also described.

The book is intended to have an applicative cut to mirror the growing interest and use of statistical methods and software in several fields of knowledge. Other than covering CTT and IRT basics, the book comprehensively accounts for the most recent theoretical IRT developments and provides `Stata` and `R` software codes for each method described within the chapters. For `Stata` we use version 12 and for `R` we use version 3.1.2. To enhance comprehension, real datasets are employed in the examples and the resulting software outputs are illustrated in detail.

The book is specifically designed for graduate students in applied statistics and psychometrics, and practitioners involved in measurement tasks within the education, health, and marketing fields, among others. For these figures, the book can be a reference work that will provide both the theoretical framework behind testing and measuring and a practical guide to carry out relevant analyses also based on the most advanced extensions and developments.

The prerequisites to properly understand the content of this book are at the level of a basic course of statistics which one can learn from the main parts of textbooks such as Mood et al. (1974) or Casella and Berger (2006). Moreover, a modest knowledge of `Stata` is required, which one can learn from the manuals for this software or from books such as Acock (2008). We use the main `Stata` module for the examples and suggest using `gllamm` (Rabe-Hesketh et al. 2004) for analyzing data from a questionnaire. Similarly, a basic knowledge of `R` is required which one may acquire from textbooks

such as Dalgaard (2008). In this regard, the main packages we use are `ltm` (Rizopoulos, 2006) and `MultiLCIRT` (Bartolucci et al., 2014), even though we acknowledge that many `R` packages exist which may be used alternatively.* Finally, the datasets used for the illustrative examples in this book as well as the main updates of `Stata` and `R` codes are available from the authors' webpage.†

The book is organized into six chapters. Chapter 1 addresses preliminary aspects related to measurement in the psychological and educational fields and to the logic, development, and use of questionnaires. The chapter also provides a description of the notation and datasets used throughout the book.

Chapter 2 examines the theory of reliability from the perspective of CTT, the procedures used to determine reliability with test data and the estimation of true scores, together with some indices commonly used for item analysis. After a brief description of the conceptual basis of test validity, the remainder of the chapter deals with test bias and with extensions of CTT (i.e., generalizability theory).

Chapter 3 is aimed at illustrating the main assumptions characterizing IRT models for dichotomously scored items and the concept of item characteristic curve. The chapter also describes the three main IRT models—that is, the well-known Rasch model, the two-parameter logistic model, and the three-parameter logistic model—entangling them in the framework of random-effects models under both the normality assumption for the latent trait and the discreteness assumption.

Chapter 4 is devoted to a generalization of IRT models for binary data to account for the categorical nature of item responses. It is very common for psychological and educational tests to be characterized by polytomous items to supply more precise information about the response process. The most relevant case we typically refer to concerns items with ordered categories. Specifically, Chapter 4 illustrates the assumptions of IRT models for polytomous items and the main criteria used to classify these models, with special attention on their statistical properties.

Chapter 5 deals with two fundamental steps of IRT analyses. The chapter's first part is devoted to estimation methods such as the joint maximum likelihood method, the conditional maximum likelihood method, and the marginal maximum likelihood method. The second part is devoted to diagnostic instruments; we distinguish: graphical tools, methods focused on the

* See `http://cran.r-project.org/web/views/Psychometrics.html` for a list of `R` packages, many of which contain functions for the methods that are used to analyze data collected by questionnaires.

† See `https://sites.google.com/site/questionnairesbook/datasets` for the datasets and `https://sites.google.com/site/questionnairesbook/software-updates` for the software updates.

goodness-of-fit measurement and based on parametric and nonparametric hypothesis tests, and methods focused on differential item functioning.

In Chapter 6, we illustrate several extensions of the models discussed previously. Traditional IRT models assume that the distribution of individual abilities, either continuous or discrete, is the same for all subjects. The first way to extend classical IRT models is to enclose individual covariates in the model so that the distribution of the ability becomes individual specific. Another way to extend classical IRT models is to allow the individual abilities to be dependent as a result of sharing common factors in the case of multilevel or longitudinal data. In addition, according to traditional IRT models, all dependencies among individual responses are accounted for by only one latent trait. However, questionnaires are often made up of several subsets of items measuring different constructs (or dimensions). Therefore, the chapter also deals with multidimensional models. It concludes with an overview of structural equation models in which the relationships between these models and IRT models are outlined.

Acknowledgments

We specially thank Roberto Ricci and Patrizia Falzetti of the Italian Institute for the Evaluation of the Education System (INVALSI) for allowing us to use some of the data they produce in our examples.

Also, we thank Guido Miccinesi of the Institute for Cancer Study and Prevention of Florence and Andrea Bonacchi of the University of Florence for allowing us to use the Hospital Anxiety and Depression Scale (HADS) data in the available examples.

Finally, we thank Phil Bardsley of the University of California and Olga Demidova and Svetlana Sokolova of the National Research University Higher School of Economics of Moscow for allowing us to use, and for their assistance with, the Russian Longitudinal Monitoring Survey (RLMS) data.

We also acknowledge that a relevant part of the research behind this book has been supported by the Grant RBFR12SHVV of the Italian Government (FIRB 2012 Project "Mixture and Latent Variable Models for Causal Inference and Analysis of Socio-Economic Data").*

*See `https://sites.google.com/site/latentfirb/`.

Contents

List of Figures