Item Response Theory Parameter Estimation Techniques

This volume guides its reader through the basics of Item Response Theory, with an emphasis on what and how to include relevant information in the methods and results sections of professional papers. The author offers examples of good and bad write-ups.

Item Response Theory clearly describes the most recently developed IRT models and furnishes detailed explanations of algorithms that can be used to estimate the item or ability parameters under various IRT models. Extensively revised and expanded, this edition offers three new chapters discussing parameter estimation with multiple groups, parameter estimation for a test with mixed item types, and Markov chain Monte

Carlo methods. It includes discussions on issues related to statistical theory, numerical methods, and the mechanics of computer programs for parameter estimation, which help to build a clear understanding of the computational demands and challenges of IRT estimation procedures. Advantages and disadvantages of joint maximum likelihood, marginal maximum likelihood, and Bayesian

methods of parameter estimation in item response theory are discussed and compared. Author supplied keywords include: Mental Test Theory and Bias (Statistical).

ASCAL: A Microcomputer Program for Estimating Logistic IRT (Item Response Theory) Item Parameters

Bayesian Item Response Modeling

Least Squares Estimation for Latent Variables with Dichotomous Item Response Data

Accuracy of Item Response Theory Parameter Estimates Using Maximum Likelihood and Bayesian Procedures as Implemented in LOGIST and BILOG

Item response theory has become an essential component in the toolkit of every researcher in the behavioral sciences. It provides a powerful means to study individual responses to a variety of stimuli, and the methodology has been extended and developed to cover many

different models of interaction. This volume presents a wide-ranging handbook to item response theory - and its applications to educational and psychological testing. It will serve as both an introduction to the subject and also as a comprehensive reference volume for practitioners and researchers. It is organized into six major sections: the nominal categories model, models for response time or multiple attempts on

items, models for multiple abilities or cognitive components, nonparametric models, models for nonmonotone items, and models with special assumptions. Each chapter in the book has been written by an expert of that particular topic, and the chapters have been carefully edited to ensure that a uniform style of notation and presentation is used throughout. As a result, all researchers whose work uses item response

theory will find this an indispensable companion to their work and it will be the subject's reference volume for many years to come. This graduate-level textbook is a tutorial for item response theory that covers both the basics of item response theory and the use of R for preparing graphical presentation in writings about the theory. Item response theory has become one of the most powerful tools used in test construction, yet

one of the barriers to learning and applying it is the considerable amount of sophisticated computational effort required to illustrate even the simplest concepts. This text provides the reader access to the basic concepts of item response theory freed of the tedious underlying calculations. It is intended for those who possess limited knowledge of educational measurement and psychometrics. Rather

than presenting the full scope of item response theory, this textbook is concise and practical and presents basic concepts without becoming enmeshed in underlying mathematical and computational complexities. Clearly written text and succinct R code allow anyone familiar with statistical concepts to explore and apply item response theory in a practical way. In addition to students of educational measurement,

this text will be valuable to measurement specialists working in testing programs at any level and who need an understanding of item response theory in order to evaluate its potential in their settings.
First thorough treatment of multidimensional item response theory
Description of methods is supported by numerous practical examples
Describes procedures for multidimensional computerized adaptive

testing

Handbook of Item Response Theory

Estimation of Parameters in Item Response Models of Psychological Testing

Principles and Applications

Item Difficulty Parameter Estimation in Item Response Theory Prior to the Administration of the Test

Volume 2: Statistical Tools

This book is open access under a CC BY-NC 2.5 license. This book describes the extensive contributions

made toward the advancement of human assessment by scientists from one of the world's leading research institutions, Educational Testing Service. The book's four major sections detail research and development in measurement and statistics, education policy analysis and evaluation, scientific psychology, and validity. Many of the developments presented have become de-facto standards in educational and psychological measurement, including in item response theory (IRT), linking and equating, differential item functioning (DIF), and educational surveys like the National Assessment of Educational Progress (NAEP), the Programme of international Student Assessment (PISA), the Progress of International Reading Literacy Study (PIRLS) and the Trends in

Mathematics and Science Study (TIMSS). In addition to its comprehensive coverage of contributions to the theory and methodology of educational and psychological measurement and statistics, the book gives significant attention to ETS work in cognitive, personality, developmental, and social psychology, and to education policy analysis and program evaluation. The chapter authors are long-standing experts who provide broad coverage and thoughtful insights that build upon decades of experience in research and best practices for measurement, evaluation, scientific psychology, and education policy analysis. Opening with a chapter on the genesis of ETS and closing with a synthesis of the enormously diverse set of contributions made over its 70-year

history, the book is a useful resource for all interested in the improvement of human assessment.

This book focuses on the practical application of statistical techniques for assessing measurement invariance with less emphasis on theoretical development or exposition. Instead, it describes the methods using a pedagogical framework followed by extensive illustrations that demonstrate how to use software to analyze real data. The chapters illustrate the practical methods to assess measurement invariance and shows how to apply them to a range of data. The computer syntax and data sets used in this book are available for download here: people.umass.edu/cswells.

In the decade of the 1970s, item response theory became the dominant

topic for study by measurement specialists. But, the genesis of item response theory (IRT) can be traced back to the mid-thirties and early forties. In fact, the term "Item Characteristic Curve," which is one of the main IRT concepts, can be attributed to Ledyard Tucker in 1946. Despite these early research efforts, interest in item response theory lay dormant until the late 1960s and took a backseat to the emerging development of strong true score theory. While true score theory developed rapidly and drew the attention of leading psychometricians, the problems and weaknesses inherent in its formulation began to raise concerns. Such problems as the lack of invariance of item parameters across examinee groups, and the inadequacy of classical test

procedures to detect item bias or to provide a sound basis for measurement in "tailored testing," gave rise to a resurgence of interest in item response theory. Impetus for the development of item response theory as we now know it was provided by Frederic M. Lord through his pioneering works (Lord, 1952; 1953a, 1953b). The progress in the fifties was painstakingly slow due to the mathematical complexity of the topic and the nonexistence of computer programs.

Estimating Parameters for Multidimensional Item Response Theory Models by MCMC Methods

On Joint and Marginal Bayesian Estimation in Item Response Theory

Parameter Estimation Techniques

A Comparison of the Effectiveness of Item Parameter Estimation

Techniques Used with the 3-parameter Logistic Item Response Theory Model Parameter Estimation in a Modified One-parameter Item Response Theory Model Using Monte-Carlo Methods

A new method for estimating the ability parameters in the statistical models based on item response theory (IRT) was presented in this study. A brief review of existing parameter estimation methods suggested the utility of a parameter estimation method which requires no specific assumption. By re-examining the performance of maximizing the ln L of the joint maximum likelihood estimation (JMLE; Lord, 1980), and by applying theorems taken from Rockafellar (1970), it was proved that ln L has a unique global maximum point. The performance of maximizing ln L and the problem caused by

maximizing ln L were analyzed mathematically. In order to avoid the problem of extreme values and to have a more plausible result, this study introduced a new objective function, $Qsb1$, to be minimized. The relationship between maximizing ln L and minimizing $Qsb1$ was also analyzed mathematically. The close-form solution of minimizing $Qsb1$ was derived algebraically. A new set of ability parameter estimators were found and called least squares estimators (LSE). The new estimator, LSE, requires no specific assumption except the general assumptions of the IRT models. The asymptotic equivalence between JMLE and LSE was briefly discussed. Some possible applications of LSE were proposed. For example, by treating the LSE's as the true values of ability parameters,

the JMLE for item parameters can be carried out. Then, two simulation studies were carried out to validate LSE and its applications. The result of the first simulation study suggested that the empirical standard error of LSE is less than the asymptotic standard error of JMLE, and the empirical bias of LSE is less than those of JMLE and Bayesian estimators. The result of the second simulation study suggested that the accuracy of LSE and its applications is comparable to that of BILOG, one of the most popular and recommended IRT parameter estimation computer programs. Some possible extensions of this study were also briefly discussed.

Item response theory (IRT) has moved beyond the confines of educational measurement into assessment

domains such as personality, psychopathology, and patient-reported outcomes. Classic and emerging IRT methods and applications that are revolutionizing psychological measurement, particularly for health assessments used to demonstrate treatment effectiveness, are reviewed in this new volume. World renowned contributors present the latest research and methodologies about these models along with their applications and related challenges. Examples using real data, some from NIH-PROMIS, show how to apply these models in actual research situations. Chapters review fundamental issues of IRT, modern estimation methods, testing assumptions, evaluating fit, item banking, scoring in multidimensional models, and advanced IRT methods.

New multidimensional models are provided along with suggestions for deciding among the family of IRT models available. Each chapter provides an introduction, describes state-of-the art research methods, demonstrates an application, and provides a summary. The book addresses the most critical IRT conceptual and statistical issues confronting researchers and advanced students in psychology, education, and medicine today. Although the chapters highlight health outcomes data the issues addressed are relevant to any content domain. The book addresses: IRT models applied to non-educational data especially patient reported outcomes Differences between cognitive and non-cognitive constructs and the challenges these bring to modeling. The application of

multidimensional IRT models designed to capture typical performance data. Cutting-edge methods for deriving a single latent dimension from multidimensional data A new model designed for the measurement of constructs that are defined on one end of a continuum such as substance abuse Scoring individuals under different multidimensional IRT models and item banking for patient-reported health outcomes How to evaluate measurement invariance, diagnose problems with response categories, and assess growth and change. Part 1 reviews fundamental topics such as assumption testing, parameter estimation, and the assessment of model and person fit. New, emerging, and classic IRT models including modeling multidimensional data and the use of new IRT models in typical

performance measurement contexts are examined in Part 2. Part 3 reviews the major applications of IRT models such as scoring, item banking for patient-reported health outcomes, evaluating measurement invariance, linking scales to a common metric, and measuring growth and change. The book concludes with a look at future IRT applications in health outcomes measurement. The book summarizes the latest advances and critiques foundational topics such a multidimensionality, assessment of fit, handling non-normality, as well as applied topics such as differential item functioning and multidimensional linking. Intended for researchers, advanced students, and practitioners in psychology, education, and medicine interested in applying IRT methods, this book also serves as a

text in advanced graduate courses on IRT or measurement. Familiarity with factor analysis, latent variables, IRT, and basic measurement theory is assumed.

By using familiar concepts from classical measurement methods and basic statistics, this book introduces the basics of item response theory (IRT) and explains the application of IRT methods to problems in test construction, identification of potentially biased test items, test equating and computerized-adaptive testing. The book also includes a thorough discussion of alternative procedures for estimating IRT parameters and concludes with an exploration of new directions in IRT research and development.

Handbook of Item Response Theory, Volume One

Theory and Applications
The Methodological, Psychological and Policy Contributions of ETS
An Extended Item Response Theory Model Incorporating Item Response Time
Item Response Theory Parameter Estimation with Response Times as Collateral Information

Information functions are used to find the optimum ability levels and maximum contributions to information for estimating item parameters in three commonly used logistic item response models. For the three and two parameter logistic models, examinees who contribute maximally to

the estimation of item difficulty contribute little to the estimation of item discrimination. This suggests that in applications that depend heavily upon the veracity of individual item parameter estimates (e.g. adaptive testing or test construction), better item calibration results may be obtained (for fixed sample sizes) from examinee calibration samples in which ability is widely dispersed. Keywords: Response theory; Information functions. (SDW).

Drawing on the work of

internationally acclaimed experts in the field, Handbook of Item Response Theory, Volume Two: Statistical Tools presents classical and modern statistical tools used in item response theory (IRT). While IRT heavily depends on the use of statistical tools for handling its models and applications, systematic introductions and reviews that emphasize their relevance to IRT are hardly found in the statistical literature. This second volume in a three-volume set fills this void. Volume Two

covers common probability distributions, the issue of models with both intentional and nuisance parameters, the use of information criteria, methods for dealing with missing data, and model identification issues. It also addresses recent developments in parameter estimation and model fit and comparison, such as Bayesian approaches, specifically Markov chain Monte Carlo (MCMC) methods.

An assessment system should be able to identify the potential of each learner

and the quality of education is dependent on providing valid score and grades to every examinee. This empirical study investigated the use of Item Response Theory (IRT) test-linking techniques (anchor-items and common-person designs) as a method of maintaining equivalent standards across years. IRT parameter estimates are assumed to be invariant. The accuracy with which each method was able to estimate item parameter estimates was also investigated.Below are the study questions,1.

How do the item parameter estimates for the two methods compare?2. How do the Item Response Functions (IRFs) and the Test Response Functions (TRFs) compare for the two designs?3. How do the Item Information Functions (IIFs) and the Test Information Functions (TIFs) compare for the two designs?4. How does the reliability and Standard Error of Measurement (SEM) compare for the two designs?The study made the following findings:- The Pearson correlation

coefficient for the item parameters are significant different.- IRFs and TRFs are significantly different for the two linking designs with the anchor-item test IRFs/TRFs approximating the theoretical Item characteristic curve (ICC).- IIFs and TIFs for the anchor-item test provide more information.- The anchor-item test is more reliable.
Fundamentals of Item Response Theory
Parameter Estimation Techniques, Second Edition
Advancing Human Assessment

Item and Person Parameter Estimation Using Hierarchical Generalized Linear Models and Polytomous Item Response Theory Models

Handbook of Item Response Theory Modeling

Item Response Theory (IRT) Models, like the one parameter, two parameters, or normal Ogive, have been discussed for many years. These models represent a rich area of investigation due to their complexity as well as the large amount of data collected

in relationship to model parameter estimation. Here we propose a new way of looking at IRT models using I-projections and duality. We use convex optimization methods to derive these models. The Kullback-Leibler divergence is used as a metric and specific constraints are proposed for the various models. With this approach, the dual problem is shown to be much easier to solve than the primal problem. In particular when there

are many constraints, we propose the application of a projection algorithm for solving these types of problems. We also consider re-framing the problem and utilizing a decomposition algorithm to solve for parameters as well. Both of these methods will be compared to the Rasch and 2-Parameter Logistic models using established computer software where estimation of model parameters are done under Maximum Likelihood Estimation

framework. We will also compare the appropriateness of these techniques on multidimensional item response data sets and propose new models with the use of I-projections. Drawing on the work of internationally acclaimed experts in the field, Handbook of Item Response Theory, Volume One: Models presents all major item response models. This first volume in a three-volume set covers many model developments that have

occurred in item response theory (IRT) during the last 20 years. It describes models for different response formats or response processes, the need of deeper parameterization due to a multilevel or hierarchical structure of the response data, and other extensions and insights. In Volume One, all chapters have a common format with each chapter focusing on one family of models or modeling approach. An introductory section in

every chapter includes some history of the model and a motivation of its relevance. Subsequent sections present the model more formally, treat the estimation of its parameters, show how to evaluate its fit to empirical data, illustrate the use of the model through an empirical example, and discuss further applications and remaining research issues.

This book develops an intuitive understanding

of IRT principles through the use of graphical displays and analogies to familiar psychological principles. It surveys contemporary IRT models, estimation methods, and computer programs. Polytomous IRT models are given central coverage since many psychological tests use rating scales. Ideal for clinical, industrial, counseling, educational, and behavioral medicine professionals and students familiar with classical testing

principles, exposure to material covered in first-year graduate statistics courses is helpful. All symbols and equations are thoroughly explained verbally and graphically.
Handbook of Modern Item Response Theory
The Comparison of Item and Person Item Response Theory (IRT) Parameter Estimates for the Anchor-Items and Common-Persons Designs
The Basics of Item Response Theory Using R
Applications to Typical Performance Assessment

Multidimensional Item Response Theory

A complete discussion of fundamental and advanced topics in Item Response Theory written by pioneers in the field In Item Response Theory, accomplished psychometricians Darrell Bock and Robert Gibbons deliver a comprehensive and up-to-date exploration of the theoretical foundations and applications of Item Response Theory (IRT). Covering both unidimensional and multidimensional IRT, as

well as related adaptive test administration of previously calibrated item banks, the book addresses the growing need for understanding of this topic as the use of IRT spreads to other fields. The first book on the topic that offers a complete and unified treatment of its subject, Item Response Theory prepares researchers and students to understand and apply IRT and multidimensional IRT to fields like education, mental health and marketing. Accessible to

first year-graduate students with a foundation in the behavioral or social sciences, basic statistics, and generalized linear models, the book walks readers through everything from the logic of IRT to cutting edge applications of the technique. Readers will also benefit from the inclusion of: • A thorough introduction to the foundations of Item Response Theory, including its logic and origins, model-based measurement, psychological scaling, and classical test theory • An

exploration of selected mathematical and statistical results, including points, point sets, and set operations, probability, sampling, and joint, conditional, and marginal probability • Discussions of unidimensional and multidimensional IRT models, including item parameter estimation with binary and polytomous data • Analysis of dimensionality, differential item functioning, and multiple group IRT Perfect for graduate students and

researchers studying and working with psychometrics in psychology, quantitative psychology, educational measurement, marketing, and statistics, Item Response Theory will also benefit researchers interested in patient reported outcomes in health research.

The modeling of item response data is governed by item response theory, also referred to as modern test theory. The eld of inquiry of item response theory has become very large and shows the enormous progress that has

been made. The mainstream literature is focused on frequentist statistical methods for - timating model parameters and evaluating model t. However, the Bayesian methodology has shown great potential, particularly for making further - provements in the statistical modeling process. The Bayesian approach has two important features that make it attractive for modeling item response data. First, it enables the possibility of incorpor- ing nondata information beyond the

observed responses into the analysis. The Bayesian methodology is also very clear about how additional information can be used. Second, the Bayesian approach comes with powerful simulation-based estimation methods. These methods make it possible to handle all kinds of priors and data-generating models. One of my motives for writing this book is to give an introduction to the Bayesian methodology for modeling and analyzing item response data. A Bayesian counterpart is presented to the many

popular item response theory books (e.g., Baker and Kim 2004; De Boeck and Wilson, 2004; Hambleton and Swaminathan, 1985; van der Linden and Hambleton, 1997) that are mainly or completely focused on frequentist methods. The usefulness of the Bayesian methodology is illustrated by discussing and applying a range of Bayesian item response models.

Drawing on the work of 75 internationally acclaimed experts in the field, Handbook of Item Response Theory, Three-Volume Set presents all major item

response models, classical and modern statistical tools used in item response theory (IRT), and major areas of applications of IRT in educational and psychological testing, medical diagnosis of patient-reported outcomes, and marketing research. It also covers CRAN packages, WinBUGS, Bilog MG, Multilog, Parscale, IRTPRO, Mplus, GLLAMM, Latent Gold, and numerous other software tools. A full update of editor Wim J. van der Linden and Ronald K. Hambleton’s

classic Handbook of Modern Item Response Theory, this handbook has been expanded from 28 chapters to 85 chapters in three volumes. The three volumes are thoroughly edited and cross-referenced, with uniform notation, format, and pedagogical principles across all chapters. Each chapter is self-contained and deals with the latest developments in IRT. Handbook of Item Response Theory, Three Volume Set The Effects of Multidimensionality on Parameter Estimation in Item Response Theory (IRT)

Item Response Theory

A Course in Item Response Theory and Modeling with Stata

Parameter Estimation Using Expected a Posteriori and Maximum Likelihood

Item Response TheoryParameter Estimation Techniques, Second EditionCRC Press

ASCAL is a microcomputer-based program for calibrating items according to the three-parameter logistic model of item response theory. ASCAL employs Lord's (1974) modified likelihood equations (for items that are omitted or not reached) and Bayesian prior distributions on the discrimination, guessing, and ability parameters to

arrive at final estimates to the item parameters. No ability parameters are produced. ASCAL uses a modified multivariate Newton-Raphson procedure for estimating item parameters. The estimation process begins by specifying starting points for the ability and item parameters. In this procedure, abilities are first estimated using the preliminary estimates of the item parameters. After ability estimates have been obtained for all examinees, they are sorted and grouped into 20 fractiles with approximately equal numbers of examinees in each. The item parameters are then estimated by assuming that the 20 fractile means are representative of the entire

ability distribution. The sequence of ability estimation, ability grouping, and parameter estimation is repeated until the item parameters converge on stable values or fail to improve. This procedure was evaluated using Monte Carlo simulation techniques.

Over the past several decades, item response theory (IRT) and item response modeling (IRM) have become increasingly popular in the behavioral, educational, social, business, marketing, clinical, and health sciences. In this book, Raykov and Marcoulides begin with a nontraditional approach to IRT and IRM that is based on their connections to classical test theory, (nonlinear) factor analysis,

generalized linear modeling, and logistic regression. Application-oriented discussions follow next. These cover the one-, two-, and three-parameter logistic models, polytomous item response models (with nominal or ordinal items), item and test information functions, instrument construction and development, hybrid models, differential item functioning, and an introduction to multidimensional IRT and IRM. The pertinent analytic and modeling capabilities of Stata are thoroughly discussed, highlighted, and illustrated on empirical examples from behavioral and social research.

Assessing Measurement Invariance for Applied Research

Maximum Likelihood and Bayesian Parameter Estimation in Item Response Theory
Improving the Accuracy of Item Response Theory Parameter Estimates Through Simultaneous Estimation and Incorporation of Ancillary Variables
Estimation of Ability Parameter in Item Response Theory (IRT) Models in the Presence of Measurement Errors
Specifying Optimum Examinees for Item Parameter Estimation in Item Response Theory
First Published in 2010. Routledge is an imprint of Taylor & Francis, an informa company.
Modeling the interaction between persons and items for binary response data, item response

theory (IRT) has been found useful in a wide variety of applications. Over the past decades, studies have been conducted on the development and application of unidimensional as well as multidimensional IRT models. However, little literature exists on IRT-based models that incorporate one general trait and several specific trait dimensions. This book, therefore, proposes such models in the Bayesian hierarchical framework, assesses their performances in various testing situations and further compares them with the conventional IRT models using Bayesian model choice techniques. Results from the analysis suggest that the proposed models offer a better way to represent the test situations not

realized in existing models. The methodology and analysis should shed some light on the development of complex IRT models and the statistical procedures for parameter estimation, and should be especially useful to professionals in educational and psychological measurement, or anyone who may be considering utilizing IRT models for assessing persons' continuous latent traits.

ABSTRACT: There is a growing need to use response time data to improve measurement quality with the increasing popularity of computerized testing. This work simultaneously models item response and response time to improve on current IRT models that do not account for response time

when there is a time limit in real testing. The joint distribution for item response and response time is presented in this work. It is specified as the product of the conditional distribution of response accuracy given response time and the marginal distribution of response time based on the lognormal distribution. A modified version of Thissen's (1983) log linear model is used to fit the response time. Marginal maximum likelihood estimation is developed and employed to estimate the item parameters. In addition, a maximum a posteriori procedure is developed and implemented to estimate person parameters. Three different simulation studies were conducted to evaluate the precision of estimation procedures. The results

of item and person parameter estimates based on MML and MAP procedures were found to be consistent and accurate.

Models

Bayesian Irt Models with General and Specific Traits

Item Response Models and Convex Optimization

Handbook of Polytomous Item Response Theory Models

Item Response Theory and Omitted Data