Psychometrika

CONTENTS

THEORY AND METHODS

NOTES AND COMMENTS

REVIEW

NOTICES

VOLUME SIXTY-NINE **JUNE 2004** **NUMBER 2**

FIRST ANNOUNCEMENT—CALL FOR PAPERS

IMPS 2005 ANNUAL MEETING OF THE PSYCHOMETRIC SOCIETY

JULY 5-8, 2005

TILBURG UNIVERSITY, THE NETHERLANDS

The 70th Annual Meeting of the Psychometric Society and the 14th International Meeting of the Psychometric Society (IMPS2005) will be held at Tilburg University, Tilburg, The Netherlands, July 5–8, 2005. Pre-conference workshops will be held Monday, July 4, 2005, with the main conference running from Tuesday to Friday, July 5–8, 2005.

Persons wishing to present talks (contributing sessions or poster sessions) should send titles and abstracts of no more than 200 words. Abstracts should be e-mailed to "pmetrika@uncg.edu." The subject header of the e-mail should include a reference to IMPS2005.

The following information should be included in attachment:
E-mail address for contact person
Name, institutional affiliation, mailing address, and e-mail address for each author
Name of the presenter of the submission
Type of submission (Contributing or Poster session)
Title of submission
Topic area of submission (see list below)
Abstract ($<$ 200 words)
The attached file should be in text format (.doc, .rtf, .txt).

References should not be included with abstracts.

Each person is permitted to present at most one contributed paper. This restriction does not prevent a keynote speaker or an invited speaker from also presenting one contributed paper, nor does it limit the number of papers on which one can be listed as a coauthor.

The deadline for abstract submission is March 1, 2005. You will be notified by March 31 whether your presentation is accepted.

Topic area choices are (AAP) Applications, (BSI) Bayesian Statistical Inference, (CDA) Categorical Data Analysis, (CTT) Classical Test Theory, (CCC) Classification, Clustering, and Correspondence Analysis, (EDA) Exploratory Data Analysis, (FAC) Factor Analysis, (GRM) Graphical Models, (IRT) Item Response Theory, (GLM) Generalized Linear Models, (LDA) Longitudinal Data Analysis, (MDS) Multidimensional Scaling, (MVA) Multivariate Analysis, (ODS) Optimal/Dual Scaling, (SEM) Structural Equation Modeling, (VCA) Variance Components Analysis, (OTR) Others (please specify)

Further details will be posted on the Society's website as they become available: http://www.psychometrika.org/meeting/2005/index.html

MANUSCRIPTS RECENTLY ACCEPTED FOR PUBLICATION

Listed in the order they were received by the Managing Editor

Selecting the Number of Classes under Latent Class Regression: A Factor Analytic Analogue
Guan-Hua Huang

Simultaneous Classification and Multidimensional Scaling with External Information
Henk A.L Kiers, Donatella Vicar, and Maurizio Vichi

Three-mode Component Analysis with Crisp or Fuzzy Partition of Units
Roberto Rocci, and Maurizio Vichi

Variability of the MAX and MIN Statistic: A Theory of the Quantile Spread as a Function of Sample Size
James Townsend and Hans Colonius

On Nonequivalence of Several Procedures of Structural Equation Modeling
Yuan, Ke-Hai Chan, Wai

Combining Speed and Accuracy to Assess Error-Free Cognitive Processes
Mark E. Glickman, Jeremy R. Gray, and Carlos J. Morales

The Applicability of Deadline Models: Comment on Glickman, Gray, and Morales (2005)
Jeff Rouder

Rejoinder: Combining Speed and Accuracy to Assess Error-Free Cognitive Processes
Mark E Glickman, Jeremy R Gray, and Carlos J. Morales

PSYCHOMETRIKA—VOL. 69, NO. 2, 167–190
JUNE 2004

GENERALIZED MULTILEVEL STRUCTURAL EQUATION MODELING

SOPHIA RABE-HESKETH

UNIVERSITY OF CALIFORNIA, BERKELEY

ANDERS SKRONDAL

NORWEGIAN INSTITUTE OF PUBLIC HEALTH, OSLO

ANDREW PICKLES

THE UNIVERSITY OF MANCHESTER

A unifying framework for generalized multilevel structural equation modeling is introduced. The models in the framework, called generalized linear latent and mixed models (GLLAMM), combine features of generalized linear mixed models (GLMM) and structural equation models (SEM) and consist of a response model and a structural model for the latent variables. The response model generalizes GLMMs to incorporate factor structures in addition to random intercepts and coefficients. As in GLMMs, the data can have an arbitrary number of levels and can be highly unbalanced with different numbers of lower-level units in the higher-level units and missing data. A wide range of response processes can be modeled including ordered and unordered categorical responses, counts, and responses of mixed types. The structural model is similar to the structural part of a SEM except that it may include latent and observed variables varying at different levels. For example, unit-level latent variables (factors or random coefficients) can be regressed on cluster-level latent variables. Special cases of this framework are explored and data from the British Social Attitudes Survey are used for illustration. Maximum likelihood estimation and empirical Bayes latent score prediction within the GLLAMM framework can be performed using adaptive quadrature in gllamm, a freely available program running in Stata.

Key words: multilevel structural equation models, generalized linear mixed models, latent variables, random effects, hierarchical models, item response theory, factor models, adaptive quadrature, empirical Bayes, GLLAMM.

Introduction

Among the milestones in the development of statistical modeling are undoubtedly the advent of comprehensive methodologies for structural equation modeling (e.g., Jöreskog, 1973) and multilevel (regression) modeling (e.g., Goldstein, 1986) and the concomitant implementation in widely available software such as LISREL (Jöreskog & Sörbom, 1989) and MLwiN (Rasbash, Browne, Goldstein, Yang, Plewis, Healy, et al., 2000). Although developed separately and for different purposes, the modeling approaches have striking similarities. Both include latent variables in the models in order to induce, and therefore explain, correlations among the responses.

Multilevel regression models are used when the data structure is hierarchical with elementary units at level 1 nested in clusters at level 2, which in turn may be nested in (super)clusters at level 3, and so on. The latent variables, or *random effects*, can be interpreted as unobserved heterogeneity at the different levels inducing dependence among all lower-level units in the same higher-level unit. Whereas random intercepts represent heterogeneity between clusters in the

---

gllamm can be downloaded from `http://www.gllamm.org`. The paper was written while Sophia Rabe-Hesketh was employed at and Anders Skrondal was visiting the Department of Biostatistics and Computing, Institute of Psychiatry, King's College London.

Requests for reprints should be sent to Sophia Rabe-Hesketh, Graduate School of Education, 3659 Tolman Hall, University of California, Berkeley, California 94720-1670, USA. E-Mail: sophiarh@berkeley.edu

0033-3123/2004-2/2001-0938-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 191–216
JUNE 2004

LOCALLY DEPENDENT LATENT TRAIT MODEL FOR POLYTOMOUS RESPONSES WITH APPLICATION TO INVENTORY OF HOSTILITY

EDWARD H. IP

WAKE FOREST UNIVERSITY

YUCHUNG J. WANG

RUTGERS UNIVERSITY

PAUL DE BOECK
MICHEL MEULDERS

K.U. LEUVEN

Psychological tests often involve item clusters that are designed to solicit responses to behavioral stimuli. The dependency between individual responses within clusters beyond that which can be explained by the underlying trait sometimes reveals structures that are of substantive interest. The paper describes two general classes of models for this type of locally dependent responses. Specifically, the models include a generalized log-linear representation and a hybrid parameterization model for polytomous data. A compact matrix notation designed to succinctly represent the system of complex multivariate polytomous responses is presented. The matrix representation creates the necessary formulation for the locally dependent kernel for polytomous item responses. Using polytomous data from an inventory of hostility, we provide illustrations as to how the locally dependent models can be used in psychological measurement.

Key words: generalized log-linear model, hybrid kernel, combination dependency models, EM algorithm, social inhibition.

1. Introduction

The item response theory (IRT) has proved to be a powerful tool in modeling individual responses to behavioral stimuli. Within an IRT framework, individual differences in a specific construct (e.g., intelligence quotient, ability, attitude, or propensity) are captured by the relative positions of the individuals' scores on a common continuous scale of the latent trait. Since the pioneering work of Lord and Novick (1968) and Fischer (1974), IRT has been extended to cover a wide range of psychological, cognitive, and educational applications. One area in which IRT has been extended is its focus on the development of locally dependent models. Two broad approaches to local dependency (LD) can be distinguished. In the first approach, LD is viewed as being unintentionally created by the *design* of a test. Subsequently, LD is treated as a nuisance factor that is peripheral to the purpose of the study. For example, in a reading comprehension test different items may share a common stem—a common reading passage. As a consequence, the usual assumptions regarding local independence—that item responses are conditionally independent given the latent trait—do not necessarily apply. While one still wants to estimate the underlying latent trait, the residual dependency not captured by the latent trait is considered a nuisance factor, which, if not taken into account in the measurement model, might lead to serious distortion. In particular, it might result in "double counting" of information contained in multiple responses that are related to a common stem.

Requests for reprints should be sent to Edward Ip, Department of Public Health Sciences, Wake Forest University School of Medicine, MRI Building, Winston-Salem, NC 27157. Email: eip@wfubmc.edu

0033-3123/2004-2/2001-0945-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 217–234
JUNE 2004

ANALYSIS OF DISTRACTOR DIFFICULTY IN MULTIPLE-CHOICE ITEMS

JAVIER REVUELTA

AUTÓNOMA UNIVERSITY DE MADRID

Two psychometric models are presented for evaluating the difficulty of the distractors in multiple-choice items. They are based on the criterion of rising distractor selection ratios, which facilitates interpretation of the subject and item parameters. Statistical inferential tools are developed in a Bayesian framework: modal a posteriori estimation by application of an EM algorithm and model evaluation by monitoring posterior predictive replications of the data matrix. An educational example with real data is included to exemplify the application of the models and compare them with the nominal categories model.

Key words: multiple-choice items, item-response models, EM algorithm, MCMC simulation, Posterior Predictive Checks.

1. Introduction

The topic of this paper is the analysis of distractor difficulty in multiple-choice items (MCI) and some of its consequences for scoring the individuals. An MCI is composed of a stem and several alternatives. One alternative is correct and the others are incorrect or distractors. The task of the subject is to select one of the alternatives.

The difficulty of a distractor depends on its attractiveness for a given population of individuals. Easy distractors can be discarded by almost all examinees, and receive few responses. On the other hand, difficult distractors have high attractiveness and response frequency.

Analysis of distractor difficulty contributes to improving the information provided by the test and identifying poorly-written items. First, the analysis of response probabilities is informative about the errors that examinees make at different ability levels. Second, distractor difficulty determines the item information function, which can be increased by a careful selection of distractors. Third, if ability is estimated by a full information procedure (based on the complete response patterns instead of dichotomous data), the data should satisfy the so-called rising selection ratios criterion (*RSRC*), which depends on distractor difficulty.

Items that fail to accomplish the *RSRC* concede more credit to individuals that select a distractor than to those selecting the correct option. Estimated ability can even decrease after the selection of the correct option. This is an awkward property, since the examinees are instructed to select the correct alternative in order to improve their estimated ability. Thus, it is necessary to develop psychometric methods to identify these items.

Two new psychometric models are introduced in this paper. One is a latent class model (denoted by *DLC*) presented to test the *RSRC*. The other is a latent trait model (denoted by *DLT*), and includes an indicator of distractor difficulty. These models are compared with the Nominal Categories Model (*NCM*; Bock, 1972, 1997), which is widely used for the analysis of MCIs and other item formats. Both the *DLT* and the *NCM* include assumptions that preclude their use in testing the *RSRC*.

This research was supported by the DGI grant BSO2002-01485.

I would like to thank Eric Maris and Vicente Ponsoda for their advice, Juan Botella for providing the data for the empirical application, and three anonymous reviewers for their comments that were essential for improving the quality of the paper.

Requests for reprints should be sent to Javier Revuelta, Departamento de Psicologia Social y Metodologia, Universidad Autónoma de Madrid, Cantoblanco 28049, Madrid, SPAIN. javier.revuelta@uam.es

0033-3123/2004-2/2001-0892-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 235–256
JUNE 2004

ASYMPTOTIC BIASES IN EXPLORATORY FACTOR ANALYSIS AND STRUCTURAL EQUATION MODELING

HARUHIKO OGASAWARA

OTARU UNIVERSITY OF COMMERCE

Formulas for the asymptotic biases of the parameter estimates in structural equation models are provided in the case of the Wishart maximum likelihood estimation for normally and nonnormally distributed variables. When multivariate normality is satisfied, considerable simplification is obtained for the models of unstandardized variables. Formulas for the models of standardized variables are also provided. Numerical examples with Monte Carlo simulations in factor analysis show the accuracy of the formulas and suggest the asymptotic robustness of the asymptotic biases with normality assumption against nonnormal data. Some relationships between the asymptotic biases and other asymptotic values are discussed.

Key words: Asymptotic biases, structural equation modeling, asymptotic variances, nonnormal data, asymptotic robustness, factor analysis.

1. Introduction

In factor analysis or, more generally, structural equation modeling, the parameters in a posited model are estimated by, for example, the Wishart maximum likelihood method. Because the parameter estimates are subject to sampling variability and estimation error, it is important to evaluate the sizes of their accuracy, which are typically summarized by the root mean square errors from the corresponding true values.

The sampling variability of the parameter estimates has been well investigated for structural models with the assumption of true models for normally and nonnormally distributed variables (Jöreskog, 1969; Jöreskog, Sörbom, du Toit, & du Toit, 1999, Section 4, Appendix A). Though these methods can also be applied, in principle, to the rotated solutions in exploratory factor analysis, somewhat tedious adaptations are required, which have been provided by Ogasawara (1998, 1999, 2000a,b), Hayashi and Yung (1999) and Yung and Hayashi (2001) based on the works of Jennrich (1973, 1974) and Archer and Jennrich (1973).

While we have these developments in stability assessment, the biases of the parameter estimates have not been well investigated in exploratory factor analysis or structural equation modeling. Shapiro (1983, Theorem 4.3) gave a general formula of the second partial derivatives of a parameter estimator with respect to sample variances and covariances for the asymptotic bias, without specifying the forms of discrepancy functions for estimation methods. Though his results are quite general and mathematically elegant, they are given without restrictions for parameters and require adaptations when we use them for actual computation by employing some estimation methods. He stated, "Theoretically the required derivatives…can be calculated in accordance with Theorem 4.3, although in practice it may be an awful task" (p. 63), where the derivatives are for the biases. Bentler and Dijkstra (1985, Equation (1.4.6)) showed the asymptotic bias of a weighted least squares estimator, which is of the same order as that of the corresponding

---

The author is indebted to the editor and anonymous reviewers for their comments, corrections, and suggestions on this paper, and to Yutaka Kano for discussion on biases.

Requests for reprints should be sent to Haruhiko Ogasawara, Department of Information and Management Science, Otaru University of Commerce, 3-5-21, Midori, Otaru 047-8501 JAPAN. E-mail: hogasa@res.otaru-uc.ac.jp

0033-3123/2004-2/2002-2011-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 257–273
JUNE 2004

ROTATION TO SIMPLE LOADINGS USING COMPONENT LOSS FUNCTIONS: THE ORTHOGONAL CASE

ROBERT I. JENNRICH

Component loss functions (CLFs) are used to generalize the quartimax criterion for orthogonal rotation in factor analysis. These replace the fourth powers of the factor loadings by an arbitrary function of the second powers. Criteria of this form were introduced by a number of authors, primarily Katz and Rohlf (1974) and Rozeboom (1991), but there has been essentially no follow-up to this work. A method so simple, natural, and general deserves to be investigated more completely. A number of theoretical results are derived including the fact that any method using a concave CLF will recover perfect simple structure whenever it exists, and there are methods that will recover Thurstone simple structure whenever it exists. Specific CLFs are identified and it is shown how to compare these using standardized plots. Numerical examples are used to illustrate and compare CLF and other methods. Sorted absolute loading plots are introduced to aid in comparing results and setting parameters for methods that require them.

Key words: Factor analysis, component loss criteria, gradient projection, hyperplane count methods, quartimax, minimum entropy, simplimax, sorted absolute loading plots, varimax.

1. Introduction

The rotation problem in factor analysis arises from a desire to find a simple and contextually meaningful relation between items and factors. Rotation methods attempt to achieve this by rotating factors to produce simple loading matrices. Unfortunately, simple loading matrices are not well defined. Thurstone has set forth a number of general principles which, vaguely stated, say a large number of small loadings and a few large ones are what one should attempt to achieve. Actually Thurstone's (1935, p. 156) conditions are precise, but in general unattainable and hence at best can only be approximated. At first attempts were made to approximate Turnstone's conditions by visually rotating hyperplanes in two-dimensional plots in an effort to maximize the number of items close to the hyperplanes. This number is called a hyperplane count. Eber (1966) attempted to implement this procedure analytically, but the hyperplane count criterion has serious discontinuities that make analytic rotation difficult. A breakthrough came when Katz and Rohlf (1974) replaced the zero-one hyperplane count for each item by a smooth function of its hyperplane distance. They considered a two-parameter family of such functions. Rozeboom (1991) introduced a more flexible four-parameter family and applied it directly to the loadings rather than to hyperplane distances. He also allowed the possibility that the function be an arbitrary growth function. We begin with this degree of generality without the growth function requirement.

More specifically, we consider a class of criteria that may be viewed as a generalization of the quartimax criterion (Newhaus & Rigley, 1954). These are defined by an arbitrary component loss function (CLF) that is evaluated at the square of each component λ_{ir} of a loading matrix Λ. The sum of these losses is the value of the corresponding CLF criterion at Λ. These criteria include the Katz and Rohlf criteria, the Rozeboom criteria, and others that will be introduced. The CLF, or what might be called the generalized hyperplane count approach, has been largely overlooked, which is unfortunate because a method so simple, natural, and general needs to be

---

The author is very indebted to a reviewer for pointing him to the generalized hyperplane count literature and to all the reviewers for valuable comments and suggestions.

Requests for reprints should be sent to Robert I. Jennrich, Dept. of Mathematics, University of California at Los Angeles, Los Angeles, CA 90095-1555, USA.

0033-3123/2004-2/2001-0925-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 275–290
JUNE 2004

INTERVAL ESTIMATION OF GAMMA FOR AN $R \times S$ TABLE

KUNG-JONG LUI

SAN DIEGO STATE UNIVERSITY

WILLIAM G. CUMBERLAND

UNIVERSITY OF CALIFORNIA, LOS ANGELES

When the underlying responses are on an ordinal scale, gamma is one of the most frequently used indices to measure the strength of association between two ordered variables. However, except for a brief mention on the use of the traditional interval estimator based on Wald's statistic, discussion of interval estimation of the gamma is limited. Because it is well known that an interval estimator using Wald's statistic is generally not likely to perform well especially when the sample size is small, the goal of this paper is to find ways to improve the finite-sample performance of this estimator. This paper develops five asymptotic interval estimators of the gamma by employing various methods that are commonly used to improve the normal approximation of the maximum likelihood estimator (MLE). Using Monte Carlo simulation, this paper notes that the coverage probability of the interval estimator using Wald's statistic can be much less than the desired confidence level, especially when the underlying gamma is large. Further, except for the extreme case, in which the underlying gamma is large and the sample size is small, the interval estimator using a logarithmic transformation together with a monotonic function proposed here not only performs well with respect to the coverage probability, but is also more efficient than all the other estimators considered here. Finally, this paper notes that applying an ad hoc adjustment procedure—whenever any observed frequency equals 0, we add 0.5 to all cells in calculation of the cell proportions—can substantially improve the traditional interval estimator. This paper includes two examples to illustrate the practical use of interval estimators considered here.

Key words: gamma, interval estimation, ordinal scale, coverage probability, maximum likelihood estimator, bootstrap method.

1. Introduction

In behavioral or psychological science, the underlying responses are often on an ordinal scale. For example, in a study of the agreement between a wife's and husband's rating of sexual fun (Hout, Duncan, & Sobel, 1987; Agresti, 1990), the responses are categorized by: never or occasionally; fairly often; very often; and almost always (Table 1). To measure the strength of the association between wife's and husband's responses, the gamma coefficient, γ, defined as the difference between the probability of concordance and the probability of discordance, given the pair of responses under comparison is untied, is probably one of the most frequently-used indices (Goodman & Kruskal, 1954; Agresti, 1984, 1990). To find an interval estimator for γ, we traditionally use Wald's confidence interval, calculated as the maximum likelihood estimator (MLE) minus or plus its estimated standard error times an appropriate percentile of a standard normal distribution (BMDP, 1988; Agresti, 1984, 1990). Since Wald's interval estimator is derived from large sample theory, it is important to evaluate its finite-sample performance. Furthermore, it is also valuable to see if there are other simple alternative interval estimators, which can outperform Wald's confidence interval.

The authors wish to thank the Associate Editor and the two referees for many valuable comments and suggestions to improve the contents and clarity of this paper. The authors also want to thank Dr. C. D. Lin for his graphic assistance.

Requests for reprints should be sent to K.-J. Lui, Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182-7720, U.S.A. E-Mail: kjl@rohan.sdsu.edu

0033-3123/2004-2/2001-0940-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 291–303
JUNE 2004

SEEING THE FISHER Z-TRANSFORMATION

CHARLES F. BOND, JR.

TEXAS CHRISTIAN UNIVERSITY

KEN RICHARDSON

TEXAS CHRISTIAN UNIVERSITY

Since 1915, statisticians have been applying Fisher's Z-transformation to Pearson product-moment correlation coefficients. We offer new geometric interpretations of this transformation.

Key words: correlation coefficient, Fisher, geometry, hyperbolic, transformation.

1. Introduction

Noting some limitations of Pearson's product-moment correlation coefficient (r), Fisher (1915) suggested a transformation

$$Z_r = \operatorname{arctanh}(r)$$

that has advantages over r. Relative to the correlation coefficient, Z_r has a simpler distribution; its variance is more nearly independent of the corresponding population parameter (Z_ρ); and it converges more quickly to normality (Johnson, Kotz, & Balakhrishnan, 1995). Fisher's Z transformation is featured in statistics texts (e.g., Casella & Berger, 2002) and is used by meta-analysts (Lipsey & Wilson, 2001).

Much has been learned about Z_r since 1915. We now know the exact distribution of Z_r for data from a bivariate normal distribution (Fisher, 1921), the exact distribution of Z_r for data from a bivariate Type A Edgeworth distribution (Gayen, 1951), and the asymptotic distribution of Z_r for virtually any data (Hawkins, 1989). We know that Z_r can be derived as a variance-stabilizing transformation or a normalizing transformation (Winterbottom, 1979). We have Taylor series expressions for the moments of Z_r and several related statistics (Hotelling, 1953).

Although scholars have been thorough in describing the analytic properties of Z_r, they have had little to say about the geometry of this transformation. True, there is a geometric flavor to certain discussions of Z_r-transformed correlation matrices (Brien, Venables, James, & Mayo, 1984). Still, the dearth of geometric knowledge about Z_r is striking, when geometric treatments of r abound (Rodgers & Nicewander, 1988).

In the current article, we offer the first geometric interpretations of Z_r to date. In section 2, we develop some Euclidean area representations. These depict r and Z_r as areas—both in the two-dimensional scatterplot and in an N-dimensional vector space. Line segments bound the areas that represent r. Line segments and hyperbolas bound the areas that represent Z_r. Our area depictions of Z_r are easy to visualize; however, the corresponding depictions of r are nonstandard. In section 3, we introduce some concepts from hyperbolic geometry that are surprisingly useful in understanding Z_r. These allow us to develop analogues to the usual geometric representations of r. There, we interpret Z_r as a slope of the least-squares regression line in a two-dimensional scatterplot and as the length of the projection of one N-dimensional vector onto

Requests for reprints should be sent to Charles Bond, `c.bond@tcu.edu`.

0033-3123/2004-2/2002-1044-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 305–315
JUNE 2004

REMARKS ON "EQUIVALENT LINEAR LOGISTIC TEST MODELS" BY BECHGER, VERSTRALEN, AND VERHELST (2002)

GERHARD H. FISCHER

UNIVERSITY OF VIENNA

This paper discusses a new form of specifying and normalizing a Linear Logistic Test Model (LLTM) as suggested by Bechger, Verstralen, and Verhelst (*Psychometrika*, 2002). It is shown that there are infinitely many ways to specify the same normalization. Moreover, the relationship between some of their results and equivalent previous results in the literature is clarified, and it is shown that the goals of estimating and testing a single element of the weight matrix, for which they propose new methods, can be reached by means of simple, well-known tools already implemented in published LLTM software.

Key words: LLTM, MIRID, identifiability.

1. The Specification of an LLTM

In a recent *Psychometrika* paper, Bechger, Verstralen, and Verhelst (2002) (henceforth abbreviated as BVV) present a new form of normalization of a Linear Logistic Test Model (LLTM), discuss equivalent ways of specifying an LLTM, construct a Lagrange Multiplier (LM) test for the H_0 that a single element of the weight matrix has a predetermined value, and develop a conditional maximum likelihood (CML) method to improve that element when the LM test has indicated a significant departure from the H_0. BVV contend that their way of defining a normalization "reveals the effect of a normalization more clearly and allows the researcher more freedom to choose a convenient normalization" (p. 124) than the traditional approach.

Although the BVV paper is formally correct and does contain new methods and ideas concerning the LLTM, the present author doubts that this way of specifying and normalizing an LLTM is advantageous. Moreover, it is surprising that BVV do not mention simpler ways to reach the goals of testing elements of the weight matrix and of estimating their value when the H_0 has been rejected; well-known tools implemented in available software products would meet the same purposes.

The usual form of specification of an LLTM has been the following (Fischer, 1983, 1995). First, a Rasch Model (RM; Rasch, 1960) is assumed to hold for the set of items in question, including local stochastic independence. Second, it is assumed that the item difficulty parameters satisfy certain linear restrictions,

$$\boldsymbol{\beta} = \mathbf{Q}\boldsymbol{\eta} + c\mathbf{1}, \tag{1}$$

where $\boldsymbol{\beta}$ is the column vector of k item (difficulty) parameters, $\mathbf{Q}$ a $k \times p$, $p < k$, matrix of given weights, $\boldsymbol{\eta}$ a column vector of p "basic" parameters measuring the effects of certain factors on item difficulty, $\mathbf{1}$ a column vector of k ones, and c an arbitrary scalar normalization constant. The elements q_{ij}, $j = 1, \ldots, p$, can be denoted the "structure" of item i.

The constant c in (1) expresses the well-known fact that in the RM the item parameters are defined only up to an arbitrary shift (to an arbitrary additive constant c). Therefore the cardinality of the set of possible normalizations is that of the real numbers. Once a particular normalization

---

Requests for reprints should be sent to Gerhard H. Fischer, Promenadegasse 21, A 1170 Wien (Vienna), AUSTRIA

0033-3123/2004-2/2002-1022-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 317–318
JUNE 2004

EQUIVALENT LLTMS: A REJOINDER

TIMO M. BECHGER, HUUB H.F.M. VERSTRALEN,
NORMAN D. VERHELST, AND GUNTER MARIS

The main purpose of our 2002 article (Bechger, Verhelst, & Verstralen, 2002) was to demonstrate that an infinite number of alternative structure matrices $\mathbf{Q}$ may be constructed that are equivalent restrictions on the item parameters. For clarity of presentation, we wanted to distinguish explicitly between the arbitrariness in the specification of the Rasch model (RM) and the arbitrariness in the specification of the structural model, which led us to specify a normalization in a manner that differs from the way preferred by Fischer. It is reassuring but hardly surprising to learn from Fischer's remarks (Fischer, 2004) that our results concerning the identifiability of the LLTM and the equivalence of different specifications are independent of the way a normalization is imposed.

Fischer's remarks mainly concern the alleged disadvantages of our way of normalizing, which are summarized in the first paragraph of section 2 of his paper. Furthermore, and this is in our view the main contribution of his paper, he demonstrates that a model where one entry in the structure matrix is considered as a parameter, can be reformulated as an LLTM. Fischer conjectures that an equivalent LLTM may also be found when there are two or more parameters in different rows of the structure matrix. In this rejoinder we will briefly respond to Fischer's critique concerning our way to specify a normalization, and illustrate by means of an example why we think that his conjecture is not true in general.

In the RM, the item parameters are defined only up to an arbitrary additive constant c. Thus, the item parameters should not be considered as a vector of k parameters but as a line b parallel to the unit vector $\mathbf{1}$ in $\mathbb{R}^k$. A normalization is a linear transformation of b to a unique point $\boldsymbol{\beta}$ of b. It is clear that there are many equivalent ways to define the same normalization but this was of no concern to us. Our aim was not to find a unique parameterization of a normalization but rather to achieve a unique parameterization of the RM. By restricting the sum $\mathbf{a}^t\mathbf{1}$ to one (Bechger, et al., 2002, p. 133), we have arbitrarily restricted the set of allowed normalizations to those that transform the line $c\mathbf{1}$ to $\mathbf{0}$. If $b = c\mathbf{1}$, the items are equally difficult and any point on b, for example, $\boldsymbol{\beta} = \mathbf{0}$, is then transformed to $\mathbf{0}$. We fail to see why this should imply that our normalization is not "universally applicable."

Applying a normalization to the LLTM gives $\boldsymbol{\beta}_a = \mathbf{L}_a\mathbf{Q}\boldsymbol{\eta} = \mathbf{Q}_a\boldsymbol{\eta}$ and it is seen that the structure matrix changes. It was emphasized in our article that this happens because going from $\mathbf{Q}\boldsymbol{\eta}$ to $\mathbf{Q}_a\boldsymbol{\eta}$ means that the original LLTM is reparameterized as a model for functions of the item difficulties that are uniquely determined under the RM; for example, differences with respect to the first item. In our view, this does clarify the nature of normalization in the LLTM. We do not consider it a disadvantage that $\mathbf{Q}_a \neq \mathbf{Q}_b$ even if $\mathbf{a}$ and $\mathbf{b}$ represent the same normalization because it illustrates the main point of our article. To wit, each of the equivalent structure matrices in section 6 of Fischer's Remarks imposes the same restriction on the item parameters and we

Requests for reprints should be sent to Timo Bechger, CITO, P.O. Box 1034, NL-6801 MG Arnhem, THE NETHERLANDS

0033-3123/2004-2/2003-1091-A $00.75/0

PSYCHOMETRIKA—VOL. 69, NO. 2, 319–321
JUNE 2004
REVIEW

REVIEWS

David Kaplan. *Structural Equation Modeling: Foundations and Extensions*, Thousand Oaks, CA: Sage Publications, 2000, 215+xviii pp., ISBN 0-7619-1407-2.

Structural equation modeling (SEM) is a statistical tool that has untapped potential for a variety of disciplines. A handful of books on this topic at different levels of complexity have emerged in the last decade: introductory (e.g., Schumacker & Lomax, 1996; Hoyle, 1995), technical (e.g., Bollen & Long, 1993; Bollen, 1989; Marcoulides & Schumacker, 1996), or applied (Cuttance & Ecob, 1987; Kline, 1998). Kaplan's book may be considered to be somewhere between the introductory and technical levels because it covers elementary issues as well as incorporates recent advancements of SEM. Kaplan has done an excellent job in integrating the voluminous modern developments of SEM into a concise handbook.

The introduction of the historical background of SEM presented in the first chapter is interesting and often not included in other similar books. The statistical foundations of SEM depicted by Kaplan are brief yet thorough. One can easily grasp the basic issues related to SEM from the reading of the first few chapters. General issues about model specification, identification, estimation, testing and interpretation are discussed in the context of path analysis, factor analysis, and general structural equation models for single and multiple groups.

The organization of these chapters, path analysis, factor analysis, and the comprehensive model, lays the foundation for the discussion of choice among model types. The fact that path analysis assumes measurement-error-free constructs should probably deserve more attention than it has been given in the literature. Although fitting the comprehensive general model may improve parameter estimates, it is often too large, in terms of model estimation, for sample sizes commonly available to social and behavioral sciences. Therefore, if the structural part of the model is of primary interest, then an alternative to estimating the complex combined model is to develop reliable and valid scales through factor analysis before estimating the structural model. Of course, psychometricians may suggest other ways of establishing reliability and validity of measurement scales.

Kaplan discusses the statistical assumptions underlying SEM at great length. He covers not only assumptions commonly considered to be relevant to SEM such as multivariate normality, large sample size, and correct model specification, but also assumptions pertaining to general statistical inference such as random sampling, and no systematic missing data. Additionally, he discusses the assumption of "exogeneity" of exogenous variables, a term that is probably unfamiliar to many readers.

Exogeneity deals with the distributional assumption of the variables. The joint distribution of the data can often be written as a product of the conditional distribution (usually the endogenous variable(s) conditioned on the exogenous variable(s)) and the marginal distribution of the variable(s) being conditioned upon. According to Kaplan, because most estimators in SEM make use of the conditional distribution, which presumably contains the parameters of interest, parameters of the marginal distribution are implicitly disregarded. Weak exogeneity is said to hold if the parameters of the marginal distribution bear no relevant information for the parameters of the conditional distribution. That is, the set of parameters of interest and the set of parameters of the marginal distribution are variation free.

Because weak exogeneity holds for normally, independently and identically distributed observations, a way to access this assumption is to access joint multivariate normality. Kaplan states

0033-3123/2004-2/2003-0004-R $00.75/0

Psychometrika
A JOURNAL OF QUANTITATIVE PSYCHOLOGY

PSYCHOMETRIC SOCIETY
MEMBERSHIP APPLICATION

☐ New Member ☐ Renewal

Name: ____________________

Affiliation: ____________________

E-Mail: ____________________ Fax: ____________________

Address where journal should be sent: ____________________

____

____

____

Payment Method (✓one): ☐ Check enclosed[†] ☐ Visa™ or ☐ Mastercard™

Card#: ____________________

Expiration Date: ________/________

Signature: ____________________

[†] Make check payable to the Psychometric Society.

Type of Membership (✓one): ☐ Regular Member ($90 USD)
☐ Family Member ($110 USD) ☐ Student Member* ($25 USD)

*Advisor's name: ____________________ *Advisor's Signature: ____________________

Return form with payment to:
PSYCHOMETRIC SOCIETY
Carol Earey
Center for Educational Research and Evaluation
University of North Carolina at Greensboro
210 Curry Building
P.O. Box 26171
Greensboro, NC 27402-6171 USA

E-Mail: cearey@aol.com
Society website: http://www.psychometricsociety.org
Journal website: http://www.psychometrika.org

Join online at http://www.psychometricsociety.org

Psychometrika

Psychometrika, the official journal of the Psychometric Society, is devoted to the development of psychology as a quantitative rational science. *Psychometrika* (ISSN 0033-3123) is published four times a year, in March, June, September, and December by the Psychometric Society, 3040 Cornwallis Road, Research Triangle Institute, P.O. Box 12194, Durham, NC 27709 U.S.A. Periodicals postage paid at Durham, NC and additional mailing offices.

POSTMASTER: Please send address changes to: Psychometrika, Psychometric Society Business Office, Center for Educational Research and Evaluation, University of North Carolina at Greensboro, 210 Curry Building, P.O. Box 26171, Greensboro, NC 27402-6171 U.S.A.

JUNE 2004, VOLUME 69, NUMBER 2

Subscription Price: The regular subscription rate is $135.00 per volume. All annual subscriptions start with the March issue and cover the calendar year. All bills and orders are payable in advance. Checks should be made payable to the *Psychometric Society*. All checks, notices of change of address, and business communications (including notification of the nonreceipt of journals) should be sent to the Business Office of the Psychometric Society. Membership in the Psychometric Society includes a subscription to *Psychometrika*. Regular members pay annual dues of $90.00, except that family members pay annual dues of $110.00. The annual dues for a student in a recognized college or university is $25.00. Credit card payments may be made on-line at http://www.psychometrika.org.

Application for membership and student membership Application for membership and student membership in the Psychometric Society, together with a check for dues for the calendar year in which application is made, should be sent to the Business Office.

Back Issues: Back issues for the previous 6 volumes only are available from the Business Office. Persons claiming replacement for issues lost in the mail must do so with the Business Office within one year. Microfilm copies of all volumes are available from Johnson Associates, Inc., P.O. Box 1017, Greenwich, CT, 06830, and Bell & Howell Information and Learning, P.O. Box 1346, Ann Arbor, MI 48106-1346 (http://wwwlib.umi.com/sim/). *Psychometrika* is abstracted in Statistical Theory and Method Abstracts (STMA), and Psychological Abstracts/PsycINFO.

Who to contact: Contact the *Psychometric Society* Business Office for subscription/membership information, contact the Editorial Office for permissions and manuscript submissions. The respective addresses appear below.

Business Office, Psychometric Society
Carol Earey, Office Manager
University of N.C. at Greensboro
210 Curry Building
P.O. Box 26171
Greensboro, NC 27402 U.S.A.
cearey@aol.com

Editorial Office, *Psychometrika*
Gwen Eastmond, Editorial Asst.
University of N.C. at Greensboro
210 Curry Building
P.O. Box 26171
Greensboro, NC 27402 U.S.A.
pmetrika@uncg.edu

(Continued on the back inside cover page)

Content: Psychometrika is devoted to the advancement of theory and methodology for behavioral data analysis in psychology, education and the social and behavioral sciences generally. It has two sections: *Theory and Methods* (T&M), and *Application Reviews and Case Studies* (ARCS). T&M articles present original research and, occasionally, insightful reviews, addressing the development of quantitative models, statistical methods, and mathematical techniques for evaluating data from psychology, the social and behavioral sciences, and related fields. ARCS articles provide benchmarks for future applications in any of these areas: *Application Reviews* can be integrative, drawing together disparate methodologies for applications, or comparative and evaluative, discussing advantages and disadvantages of one or more methodologies in applications. *Case Studies* highlight methodology that deepens understanding of substantive phenomena through more informative data analysis, and/or more elegant data description. *Psychometrika* also publishes book reviews that are of interest from a theoretical, methodological or application perspective.

Submissions: Manuscripts should be submitted electronically to the Editorial Office, using the e-mail address in the Contact Information on the front inside cover. A U.S. postal address is provided for authors without e-mail access. Prospective authors should consult the *Psychometrika* website for style and format requirements; a manuscript that fails to comply with these requirements will be returned. Submission of a paper to this journal will be taken to imply that it represents original work not previously published, and that it is not being considered elsewhere for publication. Authors may submit manuscripts to either the T&M or ARCS Section Editors, or to the Executive Editor. In all cases the Editors will make the final decision about whether the manuscript will be reviewed for T&M or ARCS. Books for review should be sent to the Book Review Editor.

Review: The manuscript review is organized by the corresponding Section Editor, often with the assistance of one of the Associate Editors. The reviews are anonymous for authors and reviewers, but not for the Associate Editor involved. Author identities will be removed as far is possible before sending out the manuscript to the reviewers.

Copyright: Authors submitting a manuscript do so with the understanding that if it is accepted for publication the copyright of the work, including the right to reproduce the work in all forms and media, shall be assigned exclusively to the Publisher. The Publisher will not refuse any reasonable request by the author for permission to reproduce any of his or her contributions to the journal.

Contact Information:

Manuscript Submissions
Editorial Office, *Psychometrika*
University of N.C. at Greensboro
210 Curry Building
P.O. Box 26171
Greensboro, NC 27402 U.S.A.
pmetrika@uncg.edu
(336) 256-0405 (fax only)

Book Reviews
Laura M. Koehly, Review Editor
Psychometrika
Department of Psychology
Texas A&M University
College Station, TX 77843-4235
lmk@psyc.tamu.edu

Society & Journal website: http//www.psychometricsociety.org or http://www.psychometrika.org

PRINTED BY CADMUS PROFESSIONAL COMMUNICATION—BYRD PRESS DIVISION
COMPOSITION BY INTEGRE TECHNICAL PUBLISHING CO., INC.