DOCUMENT RESUME

ED 401 283 TM 025 742

AUTHOR van der Linden, Wim J.
TITLE Stochastic Order in Dichotomous Item Response Models for Fixed Tests, Adaptive Tests, or Multiple Abilities. Research Report 95-02.
INSTITUTION Twente Univ., Enschede (Netherlands). Dept. of Education.
PUB DATE Dec 95
NOTE 38p.; Portions of this paper presented at the European Meeting of the Psychometric Society (9th, Leiden, Netherlands, July 4-7, 1995).
AVAILABLE FROM Bibliotheek, Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Ability; *Adaptive Testing; *Computer Assisted Testing; Foreign Countries; *Item Response Theory; *Mathematical Models; Multivariate Analysis; *Test Items
IDENTIFIERS *Dichotomous Variables; *Stochastic Analysis

ABSTRACT

Dichotomous item response theory (IRT) models can be viewed as families of stochastically ordered distributions of responses to test items. This paper explores several properties of such distributions. The focus is on the conditions under which stochastic order in families of conditional distributions is transferred to their inverse distributions, from two families of related distributions to a third family, or from multivariate conditional distributions to a marginal distribution. The main results are formulated as two theorems that apply immediately to dichotomous IRT models. One theorem holds for unidimensional models with fixed item parameters. The other theorem holds for models with multiple abilities or with random item parameters as used, for example, in adaptive testing. (Contains 2 tables and 36 references.) (Author/SLD)

TM

ED 401 283

Stochastic Order in Dichotomous Item Response Models for Fixed Tests, Adaptive Tests, or Multiple Abilities

Research Report
95-02

Wim J. van der Linden

02257442

department of EDUCATION

Division of Educational Measurement and Data Analysis

University of Twente

Stochastic Order in Dichotomous Item Response Models for Fixed Tests, Adaptive Tests, or Multiple Abilities

Wim J. van der Linden

Stochastic order in dichotomous item response model for fixed tests, adaptive tests, or multiple abilities, Wim J. van der Linden - Enschede: University of Twente, Faculty of Educational Science and Technology, December 1995. - 31 pages.

Abstract

Dichotomous IRT models can be viewed as families of stochastically ordered distributions of responses to test items. This paper explores several properties of such distributions. More in particular, it is examined under what conditions stochastic order in families of conditional distributions is transferred to their inverse distributions, from two families of related distributions to a third family, or from multivariate conditional distributions to a marginal distribution. The main results are formulated as two theorems which immediately apply to dichotomous IRT models. One theorem holds for unidimensional models with fixed item parameters. The other theorem holds for models with multiple abilities or with random item parameters as used, for example, in adaptive testing.

Stochastic Order in Dichotomous Item Response Theory

Suppose an educational or psychological test consists of a set of n dichotomously scored items indexed by i=1,...,n. Responses to item i are denoted by a random variable U_i which takes the value 1 for a correct response and the value 0 otherwise. In addition, it is assumed that examinees respond to the test items on the basis of an ability which can be represented by a (latent) unidimensional variable θ. Item response theory (IRT) offers various stochastic models to analyze the responses of examinees to the test items. Basic treatments of IRT are given, for example, in Hambleton and Swaminathan (1985) and Lord (1980).

Three different ways are available to represent an item response model. The first representation uses the idea of a <u>response function</u> to model the probabilities by which an examinee responds to an item. Let Prob$\{U_i=1 \mid \theta\}$ be the probability that an examinee with ability level θ produces a correct response to the item, and let $p_i(\theta)$ be defined as the two-parameter logistic (2-PL) function

$$p_i(\theta) \equiv [1 + \exp(-a_i(\theta - b_i))]^{-1}, \quad -\infty<\theta<\infty, \quad -\infty<b_i<\infty, \quad a_i>0, \tag{1}$$

where b_i and a_i are usually interpreted as the difficulty and discriminating power of item i, respectively. Then,

$$\text{Prob}\{U_i=1 \mid \theta\} = p_i(\theta) \equiv [1 + \exp(-a_i(\theta - b_i))]^{-1} \tag{2}$$

is an example of the response function representation of an IRT model. Alternatives to the two-parameter logistic model are the more sparsely parameterized Rasch or one-parameter logistic (1-PL) (Fischer & Molenaar, 1995) and the Birnbaum or three-parameter logistic (3-PL) model (Hambleton & Swaminathan, 1985; Lord, 1985). Throughout this paper, when we refer to IRT any of these three response models is implied. The response function representation is standard in introductory texts to IRT. This popularity is due to the fact that it allows for an immediate graphical interpretation of the values of the item parameters. For dichotomously scored responses only a function for the correct response needs to be specified; the function for the incorrect response, $1-p_i(\theta)$, is automatically fixed.

A somewhat more involved representation is based on the idea of a (parametric) family of <u>probability mass functions</u> (pmfs) for the distribution of U_i. This family can be denoted as $\{f_i(u_i \mid \theta): -\infty<\theta<\infty\}$, where

$$f_i(u_i \mid \theta) = p_i(\theta)^{u_i}[1-p_i(\theta)]^{1-u_i}, \qquad (3)$$

and $p_i(\theta)$ is defined by (1). This representation focusses on the conditional probability distribution of U_i given θ. It is standard in texts on the statistical treatment of the estimation of the values of the item and/or ability parameters. Its product over the items and examinees gives the likelihood function associated with a set of test data.

The final representation is the one of a (parametric) family of <u>cumulative distribution functions</u> (cdfs) $\{F_i(u_i \mid \theta): -\infty<\theta<\infty\}$, where

$$F_i(u_i \mid \theta) = \sum_{y=0}^{u_i} f_i(y \mid \theta), \qquad (4)$$

and $f_i(y \mid \theta)$ is given by (3). This representation is the one addressed in the current paper. In particular, the interest is in the property of stochastic order in families of cdfs as (4).

It is important to note the subtle differences between the first and the third representation. Though the logistic function itself has a well-established reputation as a cdf in certain applications, it is not used as a cdf in the first representation--let alone as a family of such functions. Second, the logistic function in (2) is monotonically increasing in θ whereas the family of cdfs in (4) is nonincreasing in θ for u_i=0,1. Though potentially confusing, these two properties of monotonicity are closely related via a well-known theorem in statistics reviewed below.

This paper shares the interests in stochastic order in response variables with several other papers which treat IRT from a nonparametric perspective. Some useful references are: Ellis and van den Wollenberg (1993); Grayson (1988); Holland (1981, 1990); Holland and Rosenbaum (1986); Huynh (1994); Junker (1991, 1993); Mokken (1971, in press); Mokken and Lewis (1982); Molenaar (in press); Ramsay (1991, in press); Rosenbaum (1984, 1985); Stout (1987, 1990); Sijtsma (1988); Sijtsma and Junker (1994); and Sijtsma and Meijer (1992). However, our point of view is fully parametric. Nevertheless, it is believed to be useful to study the consequences of certain minimal sets of assumption on response functions even if the abilities of the examinees or the properties of the items are estimated under a parametric model as in (2). This study may help to reveal certain structures in the data with otherwise might have gone unnoticed. Several examples of such structures are discussed at the end of the paper. Knowledge of such structures can, in turn, suggest new diagnostics with respect to violations of basic assumptions underlying the model.

The early work of Mokken (1971) as well as the follow up by Holland and Rosenbaum (1986) and Rosenbaum (1984, 1985) deserve special mention. These authors derived an important result for conditional covariances between item response variables from the assumptions of conditionally independent and associated items with monotonic response functions. In fact, their result is stronger than one of the results we derive. On the other hand, our interest is also in the regression of (functions of) some response variables on (functions of) other response variables as well as in generalization of the results to multidimensional response functions and tests with random item parameters as used, for example, in computerized adaptive testing. The body of this paper, however, consists of a systematic treatment of the notion of stochastic order in families of (dichotomous) (multivariate) random variables such as defined in (4). Several properties of these families will be introduced as a series of lemmas with proofs. The main results are then formulated as two theorems which follow immediately from the lemmas. One theorem holds for the conventional case of a unidimensional test with a fixed design. The other theorem specifies the conditions under which the results hold if the ability structure underlying the test is multidimensional or the test items are randomly assigned to the examinees. The final section discusses the application of the results to the analysis of data obtained through from several fixed and random test designs, including a well-known adaptive testing design.

Stochastic Order

The definition of a family of random variables stochastically ordered in a parameter is given in many textbooks (e.g., Lehmann, 1986). The same holds for the result that the expected value of a (monotonic) function of stochastically ordered variables is increasing in the parameter. A more comprehensive treatment of the notion of stochastic order typically lacks. Because of the relevance of the concept of stochastic order for dichotomous IRT, this section of the paper tries to fill the void. In particular, it examines under what conditions stochastic order is transferred to a family of inverse distributions (that is, distributions in which the random variable and parameter change their status) or from two given families of distributions to a third family. Then a few properties of stochastic order in families of multivariate (conditional) cdfs will be presented. The results will be applied to IRT in a later section.

For simplicity, the same notation will be used for all pdfs and cdfs as well as for all algebraic functions used in the treatment. Also, without explicit mention it is assumed that all pdfs and expectations exist. Finally, to avoid complications due to densities equal to zero for some values of the random variables all definitions and results are assumed to be formulated only for the supports of their pdfs.

Definition 1 (Monotone likelihood ratio). A family of (conditional) density functions $\{f(y|x)\}$ has a monotone likelihood ratio (MLR) in y w.r.t x if for any $x_1 > x_0$

$$\frac{f(y|x_1)}{f(y|x_0)}$$

does not decrease in y (e.g., Lehmann, 1986, p. 78).

Note that to obtain generality the likelihood ratio is not required to be strictly increasing in y. The same relaxation is present in the following definition.

Definition 2 (Stochastic order). A family of random variables $\{Y|x\}$ is stochastic ordered (SO) in x if, for all y, its cumulative distribution functions, $\{F(y|x)\}$, do not increase in x (e.g., Lehmann, 1986, p. 84).

As an important consequence of the fact that no strict order is required in the definition of SO, it holds that $\{Y \mid x\}$ is SO if X and Y are independent. This implication will be used when we discuss Lemma 10 below.

Observe that both definitions formalize the same idea of a random variable tending to produce larger values if another variable does the same. However, MLR is stronger than SO (see Lemma 4 below). MLR is a useful property in statistical inference whereas the assumption of SO is often made in statistical modeling because it is weaker and has the nice graphical interpretation of a family of cdfs being similarly ordered across all possible parameter values for each possible value of its argument.

It should be noted that though the property of SO seems to imply that the two variables have positive correlation, this suggestion is misleading. Positive correlation between variables would involve such properties as symmetry (positive correlation of X with Y implies correlation of Y with X), transitivity (positive correlation of X with Y and Y with Z implies positive correlation of X with Z) as well as correlation between two variables induced by a common covariate. As shown below, such properties do not hold for SO.

Expected Values

Note that if the above two definitions hold, they also hold for X and/or Y replaced by nondecreasing functions $\varphi_1(X)$ and $\varphi_2(Y)$. The well-known Lemmas 1 and 2 below are based on a multivariate version of this property.

Lemma 1. Let $\{Y_i|x;\ i=1,\ldots,n\}$ be independently distributed with densities $f(y_i|x)$ and let $\varphi(y_1,\ldots,y_n)$ be a function not decreasing in any y_i. If $\{f(y_i|x)\}$ has MLR in y_i w.r.t. x for all i, then $E[\varphi(Y_1,\ldots,Y_n)|x]$ is a nondecreasing function of x (Lehmann, 1986, p. 85, Lemma 2(i)).

Lemma 2. Under the same conditions as in Lemma 1, if $\{Y_i|x;\ i=1,\ldots,n\}$ is SO in x, then $E[\varphi(Y_1,\ldots,Y_n)|x]$ is a nondecreasing function of x (Lehmann, 1986, p. 85, Lemma 2(ii)).

Note that Lemmas 1 and 2 imply that for a single random variable Y the expected value $E[Y|x]$ is a nondecreasing function of x under the conditions given. This property is frequently used in the proofs of the lemmas presented below.

Inverse Distributions

The question can be raised under what conditions the properties of MLR and SO for a family of conditional variables, $\{Y \mid x\}$, imply MLR and/or SO for the inverse family, $\{X \mid y\}$. As it turns out, MLR is always symmetric but SO is not. However, an exception is the case of dichotomous (functions of) random variables for which the two properties coincide and symmetry of SO is implied. This case is important for the treatment of SO in dichotomous IRT models. The results are summarized as follows:

Lemma 3. $\{f(y|x_1,x_2)\}$ has MLR in y w.r.t. x_1 if and only if $\{f(x_1|y,x_2)\}$ has MLR in x_1 w.r.t. y for all x_2.

Proof. Chen, Chuang and Novick (1981, Theorem 1) offer a version of this lemma without the conditioning variable x_2. Following their argument, for any $x_1' > x_1$, if $y' > y$, then the following inequalities are equivalent:

$$\frac{f(y'|x_1',x_2)}{f(y|x_1,x_2)} \geq \frac{f(y|x_1',x_2)}{f(y|x_1,x_2)},$$

$$\frac{f(y'|x_1',x_2)}{f(y|x_1,x_2)} \geq \frac{f(y'|x_1',x_2)}{f(y|x_1,x_2)},$$

$$\frac{f(y'|x_1',x_2)}{f(y|x_1,x_2)}\frac{f(y|x_2)}{f(y'|x_2)} \geq \frac{f(y'|x_1,x_2)}{f(y|x_1,x_2)}\frac{f(y|x_2)}{f(y'|x_2)}.$$

Multiplying the left-hand and right-hand side by $f(x_1'|x_2)/f(x_1'|x_2)$ and $f(x_1|x_2)/f(x_1|x_2)$, respectively, gives:

$$\frac{f(x_1'|y',x_2)}{f(x_1'|y,x_2)} \geq \frac{f(x_1|y',x_2)}{f(x_1|y,x_2)}.$$

□

Though this property of symmetry seems to support the intuition of MLR as "positive correlation" between two variables in the sense that the events of high (low) values on two variables tend to occur simultaneously, it is easy to show by counterexample that this intuition is not valid for SO.

<u>Lemma 4</u>. If $\{f(y|x_1,x_2)\}$ has MLR in y w.r.t. x_1, then $\{Y|x_1,x_2\}$ is SO in x_1 and $\{X_1|y,x_2\}$ is SO in y.

<u>Proof</u>. From Lehmann (1986, p. 85, Lemma 2(ii)) it follows that the assumption guarantees that $\{Y|x_1,x_2\}$ is SO in x_1. The fact that $\{X_1|y,x_2\}$ is SO in y then follows from Lemma 3. □

In the following lemmas, a variable or function is called dichotomous if it can take two distinct values.

<u>Lemma 5</u>. If Y is dichotomous, then $\{f(y|x_1,x_2)\}$ has MLR in y w.r.t. x_1 if and only if $\{Y|x_1,x_2\}$ is SO in x_1.

<u>Proof</u>. Let Y have possible values y' and y, with $y'>y$. Then for any $x_1'>x_1$ the

following inequalities are equivalent:

$$\frac{f(y'|x_1', x_2)}{f(y\ |x_1, x_2)} \geq \frac{f(y|x_1', x_2)}{f(y|x_1, x_2)} ,$$

$$\frac{1-f(y|x_1', x_2)}{1-f(y|x_1, x_2)} \geq \frac{f(y|x_1', x_2)}{f(y|x_1, x_2)} ,$$

$$f(y|x_1, x_2) \geq f(y|x_1', x_2) ,$$

and

$$F(y|x_1, x_2) \geq F(y|x_1', x_2).$$

Since $F(y'|x_1,x_2)=F(y'|x_1',x_2)=1$, the required result follows. □

Lemma 6. If Y is dichotomous and $\{Y|x_1,x_2\}$ is SO in x_1, then $\{X_1|y,x_2\}$ is SO in y.

Proof. Lemmas 3 and 5. □

Lemma 7. If X is dichotomous, $\{f(y|x)\}$ has MLR in y w.r.t. x if and only if $\{X|y\}$ is SO in y.

Proof. Lemmas 3 and 5. □

Transfer of Stochastic Order

Suppose three families of conditional distributions are given which are related to each other because they share a common variable. Under what conditions does

SO for two of the families transfer to the third family?

Lemma 8. Let {Z|y} and {Y|x} be SO in y and x, respectively. Then {Z|x} is SO in x if Z and X are independent given Y=y.

Proof. It holds that

$$\begin{aligned} f(z|x) &= \int f(z,y|x)dy \\ &= \int f(z|y,x)f(y|x)dy \\ &= \int f(z|y)f(y|x)dy. \end{aligned}$$

Thus,

$$F(z|x) = \int F(z|y)f(y|x)dy.$$

F(z|y) is decreasing in y and {Y|x} is SO in x. It follows from Lemma 2 that F(z|x) decreases in x, and thus that {Z|x} is SO. □

Lemma 9. If {Y|x} and {Z|x} are SO in x, then {Z|y} is SO in y if y is dichotomous and Z and Y are independent given X=x.

Proof. Lemmas 6 and 8. □

The example in Table 1 shows that SO is not transitive. Because

[Table 1 about here]

$P(Y=1 \mid X=1)=0.30/0.50 < P(Y=1 \mid X=2)=0.20/0.50$ and $P(Z=1 \mid Y=1)=0.30/0.50 < P(Z=1 \mid Y=2)=0.25/0.50$, it follows that $F_{Y|x}(y|1) \geq F_{Y|x}(y|2)$ and $F_{Z|y}(z|1) \geq F_{Z|y}(z|2)$ for all y and z, respectively. However, $F_{Z|x}(1|1) < F_{Z|x}(1|2)$ because $P(Z=1 \mid X=1)=0.25/0.50 < P(Z=1 \mid X=2)=0.30/0.50$.

It does not hold generally that $\{Z \mid y\}$ is SO in y if $\{Y \mid x\}$ and $\{Z \mid x\}$ are. By symmetry, $\{Y|z\}$ would also be SO in z, which contradicts the earlier conclusion that SO is not symmetric. Thus the intuitive notion of two variables correlating positively if they have a "common covariate" does not apply here either.

Multivariate Conditioning Variables

Families of conditional distributions with more than one conditioning variable are introduced and the question is raised if the property of SO is maintained if the transition to a single conditioning variable is made. The question is relevant for the treatment of stochastic order in IRT models for multivariate abilities or when an item parameter becomes stochastic and the model implies stochastic order w.r.t. this parameter as well. For simplicity, only the case of two conditioning variables is discussed but generalization to larger numbers of conditioning variables is readily obtained.

The family $\{Y \mid x_1,x_2\}$ is defined to be SO in x_1 and x_2 if $F(y \mid x_1,x_2)$ is nondecreasing in x_1 for all x_2 and in x_2 for all x_1. The following lemma identifies a condition under which SO is transferred to $\{Y \mid x_1\}$:

Lemma 10. Let Y be a continuous random variable with density function f(y). Further, $\{Y|x_1,x_2\}$ is assumed to be SO in x_1 and x_2. Then $\{Y|x_1\}$ is SO in x_1 if $\{X_2|x_1\}$ is SO in x_1.

Proof. The lemma is proved as follows:

$$f(y|x_1) = \int f(y,x_2|x_1)dx_2$$
$$= \int f(y|x_1,x_2)f(x_2|x_1)dx_2.$$

Thus,

$$F(y|x_1) = \int F(y|x_1,x_2)f(x_2|x_1)dx_2.$$

Since $F(y|x_1,x_2)$ is decreasing in x_2 and $\{X_2|x_1\}$ is SO in x_1, it follows from Lemma 2 that $\{Y|x_1\}$ is SO in x_1. □

Note that the fact that condition of $\{X_2 \mid x_1\}$ being SO in x_1 implies that Lemma 10 holds if X_1 and X_2 are independent. A direct proof of this implication can be based on the fact that under independence $F(y|x_1) = \int F(y|x_1,x_2)f(x_2|x_1)dx_2 = \int F(y|x_1,x_2)f(x_2)dx_2$. Since $F(y|x_1,x_2)$ and $f(y)$ are continuous, $f(y)$ does not change sign, and $\int f(y)dy$ converges by definition, the weighted mean-value theorem for integrals (Apostol, 1967, sect. 3.19) shows that there exist a constant c such that $F(y|x_1) = F(y|x_1,c)\int f(x_2)dx_2 = F(y|x_1,c)$, which, by assumption, is decreasing in x_1.

The lemma thus shows that to proceed from a multivariate to a marginal condition, the multivariate condition has to demonstrate SO itself. The lemma is also given in van der Linden and Vos (in press). Note that for X_1 and X_2 being independent, Lemma 8 is a special case of Lemma 10.

Multivariate Distributions

A multivariate family of random variables $\{Y_1,\ldots,Y_n \mid x\}$ is defined to be SO in x if $\{F(y_1,\ldots,y_n \mid x)\}$ does not increase in x for all $(y_1,\ldots,y_n)$.

The example in Table 2 shows that SO in a series of families of univariate

[Table 2 about here]

distribution functions does not imply multivariate SO. For example, $F_{Y_1|x}(0|0) = (0.25+0.15)/0.50 > F_{Y_1|x}(0|1) = (0.10+0.00)/0.50$. The same relation holds for $F_{Y_2|x}(0|x)$. However, $F_{Y_1,Y_2|x}(1,0|0) = 0.05/0.50 < F_{Y_1,Y_2|x}(1,0|1) = 0.10/0.50$.

The following lemma identifies a condition under which multivariate SO does follow from univariate SO:

Lemma 11. If each $\{Y_i|x\}$, $i=1,\ldots,n$, is SO in x, then $\{Y_1,\ldots,Y_n|x\}$ is SO in x if $\{Y_i|x\}$, $i=1,\ldots,n$, are independent.

Proof. The lemma follows immediately from the fact that the univariate cdfs are nonnegative and do not increase in x. □

The reverse implication, however, does hold generally:

Lemma 12. If $\{Y_1,\ldots,Y_n \mid x\}$ is SO in x, then any subset of variables is SO in x.

Proof. A proof will be given for the case of two variables. For any $x' > x$,

$$F(y_1,y_2|x') \le F(y_1,y_2|x)$$

for all values of (y_1, y_2). Thus,

$$F(y_1|x') = \lim_{y_2 \to \infty} F(y_1, y_2|x')$$
$$\leq \lim_{y_2 \to \infty} F(y_1, y_2|x)$$
$$= F(y_1|x)$$

for all values of y_1. □

Functions of Random Variables

The following lemma summarizes several results for (multivariate) functions of random variables with the property of stochastic order in a common conditioning variable:

Lemma 13. Let $\{Y_i|x\}$, $i=1,\ldots,n$, be independent and SO in x, and let $\varphi_1 = \varphi_1(y_1,\ldots,y_p)$, $\varphi_2 = \varphi_2(y_{p+1},\ldots,y_q)$ and $\varphi_3 = \varphi_3(y_{q+1},\ldots,y_n)$, $0<p<q<n$, be nondecreasing in each of their arguments. If φ_3 is (1) dichotomous or a (2) nondecreasing function of $\Sigma_{i=q+1}^{n} y_i$ with each y_i dichotomous, it holds that:

(1) $\{\phi_1, \phi_2, \phi_3|x\}$ is SO in x;
(2) $\{\phi_1, \phi_2|\varphi_3\}$ is SO in φ_3;
(3) $\{\phi_j|\varphi_3\}$, $j=1,2$, is SO in φ_3;
(4) $\{\phi_j|x, \varphi_k\}$, $j \neq k=1,\ldots,3$, is SO in x for all values of φ_k;
(5) $\{X|\varphi_j, \varphi_3\}$, $j=1,2$, is SO in φ_3 for all values of φ_j;
(6) $\{\phi_j|\varphi_k, \varphi_3\}$, $j \neq k=1,2$, is SO in φ_3 for all values of φ_k.

Proof. The parts of the lemma are proved as follows:

(1) $\{\phi_j|x\}$, $j=1,\ldots,3$, are independent and SO in x. Hence, Lemma 11 gives the required result.

(2) For the cdf of the joint conditional distribution, it holds that

$$\begin{aligned} F(\varphi_1,\varphi_2|\varphi_3) &= \int F(\varphi_1,\varphi_2|\varphi_3,x)f(x|\varphi_3)dx \\ &= \int F(\varphi_1,\varphi_2|x)f(x|\varphi_3)dx. \end{aligned}$$

From Lemmas 13(1) and 12 it follows that $F(\varphi_1,\varphi_2|x)$ does not increase in x and that $\{\phi_3|x\}$ is SO in x. If φ_3 is dichotomous, Lemma 6 shows that $\{X|\varphi_3\}$ is SO in φ_3. If φ_3 is an nondecreasing function of $\sum_{i=q+1}^{n} y_i$, it follows from Lemma 4 together with Grayson's (1988; see also Huynh, 1994) result of MLR for the family of density functions associated with $\Sigma_{i=q+1}^{n} Y_i|x$ that $\{X|\varphi_3\}$ is also SO in x. In either case, Lemma 2 gives us the desired result.

(3) Lemmas 13(2) and 12.

(4) As ϕ_j and ϕ_k are independent given x,

$$F(\varphi_j|x,\varphi_k) = F(\varphi_j|x),$$

and the result follows immediately.

(5) Lemmas 13(4) and 6.

(6) It holds that

$$\begin{aligned} F(\varphi_j|\varphi_k,\varphi_3) &= \int F(\varphi_j,x|\varphi_k,\varphi_3)dx \\ &= \int F(\varphi_j|x)f(x|\varphi_k,\varphi_3)dx. \end{aligned}$$

By assumption $F(\varphi_j|x)$ is not increasing in x whereas Lemma 13(5) shows that $\{X|\varphi_k,\varphi_3\}$ is SO in φ_3 for all values of φ_k. Thus, Lemma 2 gives the desired result. □

Main Theorems

The main results for the conditional expectations and covariances between functions on variables $\{Y_i \mid x\}$ and $\{Y_i \mid x_1, x_2\}$, i=1,...,n, are formulated in the following two theorems.

Theorem 1. Let $\{Y_i|x\}$, i=1,...,n, be independent and SO in x, and let $\varphi_1 = \varphi_1(y_1,...,y_p)$, $\varphi_2 = \varphi_2(y_{p+1},...,y_q)$ and $\varphi_3 = \varphi_3(y_{q+1},...,y_n)$, $0<p<q<n$, be nondecreasing in each of their arguments. If φ_3 is (1) dichotomous or a (2) nondecreasing function of $\Sigma_{i=q+1}^{n} y_i$ with each y_i dichotomous, then:

(1) $E(\phi_j|\varphi_3)$, j=1,2, is a nondecreasing function of φ_3;

(2) $Cov(\phi_1,\phi_2|\varphi_3) \geq 0$;

(3) $Cov(\phi_j,\phi_k) \geq 0$, j,k=1,...,3; $j \neq k$.

Proof. The three parts of the theorem are proved as follows:

(1) Lemmas 13(3) and 2.

(2) Note that

$$Cov(\varphi_1,\phi_2|\varphi_3) = Cov((\phi_1,E(\phi_2|\phi_1)|\varphi_3)).$$

Let $\tau(\varphi_1,\varphi_3) \equiv E(\phi_2|\varphi_1,\varphi_3)$. Lemmas 13(6) and 2 show that τ is a nondecreasing function of φ_3. It is now to be proved that

$$\begin{aligned}\mathrm{Cov}(\phi_1,\tau(\phi_1,\varphi_3)\,|\,\varphi_3) &= E(\phi_1\,\tau(\phi_1,\varphi_3)\,|\,\varphi_3] - E(\phi_1\,|\,\varphi_3)E(\tau(\phi_1,\varphi_3)\,|\,\varphi_3)\\ &= E[(\phi_1 - E(\phi_1))\tau(\phi_1,\varphi_3)\,|\,\varphi_3]\\ &\geq 0.\end{aligned}$$

Following an argument in Casella and Berger (1990, sect. 4.7.2),

$$\begin{aligned}&E[(\phi_1 - E(\phi_1))\tau(\phi_1,\varphi_3)\,|\,\varphi_3]\\ &= E[(\phi_1 - E(\phi_1))\tau(\phi_1,\varphi_3)I_{(-\infty,0)}(\phi_1 - E(\phi_1))\,|\,\varphi_3]\\ &\quad + E[(\phi_1 - E(\phi_1))\tau(\phi_1,\varphi_3)I_{[0,\infty)}(\phi_1 - E(\phi_1))\,|\,\varphi_3]\\ &\geq E[(\phi_1 - E(\phi_1))\tau(E(\phi_1\,|\,\varphi_3),\varphi_3)I_{(-\infty,0)}(\phi_1 - E(\phi_1))\,|\,\varphi_3]\\ &\quad + E[\phi_1 - E(\phi_1))\tau(E(\phi_1\,|\,\varphi_3),\varphi_3)I_{[0,\infty)}(\phi_1 - E(\phi_1))\,|\,\varphi_3]\\ &= \tau(E(\phi_1\,|\,\varphi_3),\varphi_3)E(\phi_1 - E(\phi_1)\,|\,\varphi_3)\\ &= 0.\end{aligned}$$

(3) It holds that

$$\mathrm{Cov}(\phi_j,\phi_k) = E(\mathrm{Cov}(\phi_j,\phi_k\,|\,X)) + \mathrm{Cov}(E(\phi_j\,|\,X),E(\phi_k\,|\,X)).$$

From the previous part of the theorem it follows that the first term is the expected value of a nonnegative statistic. As $E(\phi_j\,|\,X)$ and $E(\phi_k\,|\,X)$ are nondecreasing in X, a repetition of the argument in the previous part of this proof shows that the second term is nonnegative. □

It is important to observe that all three implications in Theorem 1 address properties of regression and covariance functions which can be observed in large samples. We will return to this point in the last section when applications to IRT are discussed more directly. Individual parts of the theorem can be found in other places in the psychometric literature. However, they were established using different methods of proof than the one based on the set of lemmas derived above. The covariance property in the third part of the theorem was given earlier in Mokken (1971) and Holland (1981). Esary, Proschan, and Walkup (1967) define the covariance property in the third part of the theorem as association between the underlying sets of random variables but gave no conditions under which the property of association holds. Ahmed, León, and Proschan (1981, sect. 3.5) derive association between ϕ_1 and ϕ_2 under the same conditions as used here. Ellis (1993) proofs the third part of the theorem to be valid for any subpopulation of examinees. An important reference is Rosenbaum (1984) who gives a version of the second part of the theorem not based on the assumption on φ_3 made here. Finally, Junker (1993) establishes the first part of the theorem as the property of manifest homogeneity.

All assumptions of Theorem 1 are assumed to hold in the next theorem. The use of double indices is only to refer to rows and columns in an item x person matrix with response data. The critical event in the theorem is the presence of more than one parameters needed to characterize the distributions of the variables.

Theorem 2. Let $\{Y_{ij}|x_1,x_2\}$, i=1,...,n, j=1,...,m, be independent and SO in x_1 and x_2. P and Q are defined to be the sets of indices of two disjoint subsets of variables of $\{Y_{ij}$;1=1,...,n, j=1,...,m}. Let $\varphi_P=\varphi_P(.)$ and $\varphi_Q=\varphi_Q(.)$ be two functions nondecreasing in each of the variables with indices in P and Q, respectively. It is assumed that $\varphi_Q(.)$ is either dichotomous or nondecreasing in $\Sigma_{(i,j)\in Q}\, y_{ij}$ with each y_{ij} dichotomous. Finally, $\{X_2|x_1\}$ and $\{X_1|x_2\}$ are assumed to be SO in x_1 and x_2, respectively. It holds that:

(1) $E(\phi_K|x_v)$, K=P,Q, is a nondecreasing function of x_v, v=0,1;

(2) $E(\phi_P | x_v, \varphi_Q)$ is a nondecreasing function of x_v, $v=1,2$, for all values of φ_Q.

Proof. The two parts of this theorem follow immediately from the previous lemmas:

(1) Lemmas 13(1), 10, 12, and 2.

(2) Theorem 2(1) and Lemma 13(4) (conditional independence).

This theorem identifies the conditions under which order with respect to one parameter is maintained in a response model with more than one random person or a person and an item parameter. For example, the theorem implies that the expected sum of scores in any part of the data matrix is ordered in either parameter provided the parameters are independent or stochastically ordered themselves. The same feature holds for the expected column and row sums of the data matrix. The theorem thus reveals the conditions under which the row and column sums are ordered by a person and item parameter. Other consequences from the two theorems are presented in more detail in the corollaries in the next section.

Applications to IRT

As explained in the introduction, a dichotomous IRT model can be represented by a family of cdfs $\{F(u_i|\theta: -\infty<\theta<\infty\}$ fully determined by the probabilities $\{f_i(1|\theta): -\infty<\theta<\infty\}$ modeled as a (strictly) increasing function of θ (Lemma 6). Since $\{F(u_i|\theta)\}$ is (strictly) decreasing in θ, this family is SO in θ. Also, because the response variables U_i are dichotomous, it holds that $\{f(u_i|\theta)\}$ has MLR in u_i

w.r.t. θ. Finally, the usual assumption of local independence between response variables for different items guarantees the conditional independence required in some of the lemmas and the two theorems above.

As already observed, both theorems involve several properties which can be observed in large samples of test data. The most important properties implied by Theorem 1 are summarized in Corollary 1. Some of these properties have also been listed elsewhere (see, for example, Rosenbaum, 1984, or Sijtsma & Junker, 1994).

Corollary 1. For any dichotomous IRT model with a single ability parameter and a fixed test design it holds that:

(1) conditional item π-values given the (number-right) score on another item or subtest are nondecreasing functions of the conditioning score;

(2) item-rest regression, defined as the regression of an item score on the (number-right) score on the remaining items, is a nondecreasing function of the latter;

(3) the probability of passing a cutoff score on a subtest is a nondecreasing function of the number-right score on another subtest;

(4) if π_i^H and π_i^L are the π-values of item i in a high-scoring and low-scoring subpopulation, respectively, it holds that $D \equiv \pi_i^H - \pi_i^L$ is nonnegative;

(5) all correlations between item score are nonnegative;

(6) all item-rest correlations (item discrimination indices) are nonnegative;

(7) all previous properties hold in any subpopulation defined by number-right scores on other items or subtests;

(8) all previous properties hold for weighted scores, provided the weights are

nonnegative.

Several of these properties do already have a long tradition as a criterion for item selection in classical item analysis. For example, attempts to maximize the internal consistency of a test have always been directed at removing items with negative intercorrelations and/or item-rest correlations from the test. In addition, the corollary confirms the status of D, typically defined using Kelley's (1939) 27% rule, as a quick alternative to the item discrimination index which was popular in the pre-computer era. The notion that so-called formula scoring can be treated as equivalent to simple number-right scoring is another intuitive notion given a mathematical basis by the corollary. The corollary finally implies that classical item analysis is an effective first step to weed out items not fitting a dichotomous IRT model.

Corollary 2. In an IRT model, the properties of SO hold for a single item difficulty or ability parameter if: (1) the values of the item difficulty parameter are fixed; or (2) the values of the item difficulty parameters are random but all items are administered to the same examinees.

In both test designs, the ability of the examinees and the item difficulty parameter are independent. Since independence implies that the distribution of one parameter is (not strictly) SO in the other, Theorem 2 holds. An example of the second design is a test sampled at random from an item pool and then administered to all of the examinees in the sample (domain-referenced testing).

On the other hand, in adaptive testing, the assumption of independence between the parameters is unlikely to hold since adaptive procedures invariable use item selection rules in which more able examinees tend to get more difficult items. This feature, however, suggests that the use of such rules may lead to distributions of the values of the difficulty parameter which are SO in the ability parameter. Suppose no constraints on the availability of the values of the item difficulty parameter exist in the item pool. The following corollary shows that the results in Theorem 2 apply to a currently popular procedure of adaptive testing:

Corollary 3. For an adaptive test from a 1-PL item pool based on the maximum information principle in combination with EAP estimation of ability, the distribution of any monotonically nondecreasing function of the examinee's response vector is SO in θ.

The following argument explains the corollary. Let $\varepsilon_k \equiv -b_k$ be the value of the easiness parameter of the kth item in the adaptive test. Then $\varepsilon_k = \varepsilon_k(u_1,...,u_{k-1}) = E(\theta | u_1,...,u_{k-1})$. However, since $\{\theta | u_1,...,u_{k-1}\}$ is SO in $u_1,...,u_{k-1}$ (Lemma 6), it follows that $\varepsilon_k(u_1,...,u_{k-1})$ is nondecreasing in each of its arguments. Because $\{U_1,...,U_{k-1} | \theta\}$ is SO in θ, it follows that $\{\varepsilon_k(U_1,...,U_{k-1}) | \theta\}$ is SO in θ (Lemma 2). Note that Lemma 6 holds for any prior $f(\theta)$. In a fully Bayesian procedure, the prior can thus be chosen to be independent of the one for the item parameter to allow us to ignore the items in the pool not used in the test (for this condition of independence, see Mislevy & Wu, 1988).

The following corollary summarizes a result for IRT models with a two-dimensional ability structure:

Corollary 4. The result in Theorem 2 holds marginally in θ_1 if $\{\Theta_1 | \theta_2\}$ is a location family with conditional pdfs $f(\theta_1 - \mu | \theta_2)$.

The fact that location families have the property of SO is well documented (e.g., Lehmann, 1986, p. 84-85). An important application is the case of a bivariate normal ability distribution with constant conditional variances. As the values of the ability parameters are not controlled by design, it is a matter of empirical fact whether or not the condition in this corollary holds satisfactorily in practice. A statistical test for this condition could be based on the class of models with multivariate ability presented in Glas (1992).

References

Ahmed, A.N, León, R., & Proschan, F. (1981). Generalized association, with applications in multivariate statistics. Annals of Statistics, 9, 168-176.

Apostol, T.M. (1967). One-variable calculus, with an introduction to linear algebra (Vol. 1). New York, NY: Wiley.

Casella, G., & Berger, R.L. (1990). Statistical inference. Pacific Grove, CA: Wadsworth & Brooks/Cole.

Ellis, J.L. (1993). Subpopulation invariance of patterns in covariance matrices. British Journal of Mathematical and Statistical Psychology, 46, 231-254.

Ellis, J.L., & van den Wollenberg, A.L. (1993). Local homogeneity in latent trait models: A characterization of the homogenous monotone IRT model. Psychometrika, 58, 417-429.

Esary, J.D., Proschan, F., & Walkup, D.W. (1967). Association of random variables, with applications. Annals of Mathematical Statistics, 38, 1466-1474.

Fischer, G.H. & Molenaar, I.W. (Eds.) (1995). Rasch models: Foundations, recent developments, and applications. New York City, NY: Springer-Verlag.

Glas, C.A.W. (1992). A Rasch model with a multivariate distribution of ability. In M. Wilson (Ed.), Objective measurement: Theory into practice (Vol. 1). Norwood, NJ: Ablex.

Grayson, D.A. (1988). Two-group classification in latent trait theory: Scores with monotone likelihood ratio. Psychometrika, 53, 283-292.

Hambleton, R.K., & Swaminathan, H. (1985). Item response theory: Principles and applications. Boston, MA: Kluwer-Nijhoff Publishing.

Holland, P.W. (1981). When are item response models consistent with observed data? Psychometrika, 46, 79-92.

Holland, P.W. (1990). On the sampling theory foundations of item response theory models. Psychometrika, 55, 577-601.

Holland, P.W., & Rosenbaum, P.R. (1986). Conditional association and unidimensionality in monotone latent variable models. Annals of Statistics. 14, 1523-1543.

Huynh, H. (1994). A new proof for monotone likelihood ratio for the sum of independent Bernoulli random variables. Psychometrika, 59, 77-79.

Junker, B.W. (1991). Essential independence and likelihood-based ability estimations for polytomous items. Psychometrika, 56, 255-278.

Junker, B.W. (1993). Conditional association, essential independence, and monotone unidimensional item response models. The Annals of Statistics, 21, 1359-1378.

Kelley, T.L. (1939). The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, 30, 17-24.

Lehmann, E.L. (1986). Testing statistical hypotheses (2nd Ed.). New York, NY: Wiley.

Lord, F.M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Lawrence Erlbaum.

Mislevy, R.J., & Wu, P.-K. (1988). Inferring examinee ability when some item responses are missing (Research Report 88-8). Princeton, NJ: Educational Testing Service.

Mokken, R.J. (1971). A theory and procedure of scale analysis. Den Haag, The Netherlands: Mouton.

Mokken, R.J. (in press). Nonparametric models for dichotomous responses. In W.J. van der Linden & R.K. Hambleton (Eds.), Handbook of modern item response theory. New York, NY: Springer-Verlag.

Mokken, R.J., & Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. Applied Psychological Measurement, 6, 417-430.

Molenaar, W. (in press). Nonparametric models for polytomous responses. In W.J. van der Linden & R.K. Hambleton (Eds.), Handbook of modern item response theory. New York, NY: Springer-Verlag.

Rosenbaum, P.R. (1984). Testing the conditional independence and monotonicity assumptions of item response theory. Psychometrika, 49, 425-435.

Rosenbaum, P.R. (1985). Comparing distributions of item responses for two groups. British Journal of Mathematical and Statistical Psychology, 38, 206-215.

Ramsay, J.O. (1991). Kernel smoothing approaches to nonparametric item characteristic curve estimation. Psychometrika, 56, 611-630.

Ramsay, J.O. (in press). A functional approach to modeling test data. In W.J. van der Linden & R.K. Hambleton (Eds.), Handbook of modern item response theory. New York, NY: Springer-Verlag.

Stout, W.F. (1987). A nonparametric approach for assessing latent trait unidimensionality. Psychometrika, 52, 589-617.

Stout, W.F. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. Psychometrika, 55, 293-325.

Sijtsma, K. Contributions to Mokken's nonparametric item response theory. Unpublished doctoral dissertation, University of Groningen, The Netherlands.

Sijtsma, K., & Junker, B.W. (1994). A survey of theory and methods of invariant item ordering (Methods Series MS-94-4). Utrecht, The Netherlands: Department of Methodology and Statistics, Faculty of Social Sciences, Utrecht University.

Sijtsma, K., & Meijer, R.R. (1992). A method for investigating the intersection of item response functions in Mokken's nonparametric IRT model. Applied Psychological Measurement, 16, 149-158.

Stout, W.F. (1987). A nonparametric approach for assessing latent trait unidimensionality. Psychometrika, 52, 589-617.

Stout, W.F. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. Psychometrika, 55, 293-325.

van der Linden, W.J., & Vos, H.J. (in press). A compensatory approach to optimal selection with mastery scores. Psychometrika, 61.

Author Note

Portions of this paper were presented at the 9th European Meeting of the Psychometric Society, Leiden, The Netherlands, July 4-7, 1995. The author is indebted to Sebie J. Oosterloo, Hans J. Vos, and Rob R. Meijer for their comments on a previous version of the manuscript.

Table 1

Numerical example showing that SO is not transitive

	X=1		X=2	
	Y=1	Y=2	Y=1	Y=2
Z=1	.20	.05	.10	.20
Z=2	.10	.15	.10	.10

Table 2

Numerical example showing that univariate

SO does not imply multivariate SO

	X=1		X=2	
	$Y_1=1$	$Y_2=2$	$Y_1=1$	$Y_2=2$
$Y_2=1$	.25	.05	.10	.10
$Y_2=2$	.15	.15	.00	.30

Titles of Recent Research Reports from the Department of Educational Measurement and Data Analysis.
University of Twente, Enschede,
The Netherlands.

RR-95-02 W.J. van der Linden, *Stochastic Order in Dichotomous Item Response Models for Fixed Tests, Adaptive Tests, or Multiple Abilities*

RR-95-01 W.J. van der Linden, *Some decision theory for course placement*

RR-94-17 H.J. Vos, *A compensatory model for simultaneously setting cutting scores for selection-placement-mastery decisions*

RR-94-16 H.J. Vos, *Applications of Bayesian decision theory to intelligent tutoring systems*

RR-94-15 H.J. Vos, *An intelligent tutoring system for classifying students into Instructional treatments with mastery scores*

RR-94-13 W.J.J. Veerkamp & M.P.F. Berger, *A simple and fast item selection procedure for adaptive testing*

RR-94-12 R.R. Meijer, *Nonparametric and group-based person-fit statistics: A validity study and an empirical example*

RR-94-10 W.J. van der Linden & M.A. Zwarts, *Robustness of judgments in evaluation research*

RR-94-9 L.M.W. Akkermans, *Monte Carlo estimation of the conditional Rasch model*

RR-94-8 R.R. Meijer & K. Sijtsma, *Detection of aberrant item score patterns: A review of recent developments*

RR-94-7 W.J. van der Linden & R.M. Luecht, *An optimization model for test assembly to match observed-score distributions*

RR-94-6 W.J.J. Veerkamp & M.P.F. Berger, *Some new item selection criteria for adaptive testing*

RR-94-5 R.R. Meijer, K. Sijtsma & I.W. Molenaar, *Reliability estimation for single dichotomous items*

RR-94-4 M.P.F. Berger & W.J.J. Veerkamp, *A review of selection methods for optimal design*

RR-94-3 W.J. van der Linden, *A conceptual analysis of standard setting in large-scale assessments*

RR-94-2 W.J. van der Linden & H.J. Vos, *A compensatory approach to optimal selection with mastery scores*

RR-94-1 R.R. Meijer, *The influence of the presence of deviant item score patterns on the power of a person-fit statistic*

RR-93-1 P. Westers & H. Kelderman, *Generalizations of the Solution-Error Response-Error Model*

RR-91-1 H. Kelderman, *Computing Maximum Likelihood Estimates of Loglinear Models from Marginal Sums with Special Attention to Loglinear Item Response Theory*

RR-90-8 M.P.F. Berger & D.L. Knol, *On the Assessment of Dimensionality in Multidimensional Item Response Theory Models*

RR-90-7 E. Boekkooi-Timminga, *A Method for Designing IRT-based Item Banks*

RR-90-6 J.J. Adema, *The Construction of Weakly Parallel Tests by Mathematical Programming*

RR-90-5 J.J. Adema, *A Revised Simplex Method for Test Construction Problems*

RR-90-4 J.J. Adema, *Methods and Models for the Construction of Weakly Parallel Tests*

RR-90-2 H. Tobi, *Item Response Theory at subject- and group-level*

RR-90-1 P. Westers & H. Kelderman, *Differential item functioning in multiple choice items*

Research Reports can be obtained at costs from Bibliotheek, Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.

department of
EDUCATION

A publication by
the Department of Education
of ~~the University of Twente~~
P.O. Box 217
7500 AE Enschede
The Netherlands

NOTICE

REPRODUCTION BASIS

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").